Detecting Vietnam War Bomb Craters in Declassified Historical KH-9 Satellite Imagery

Philipp Barthelme[1], Eoghan Darbyshire[2], Dominick Vincent Spracklen[3], and Gary R Watmough[1]

[1]University of Edinburgh
[2]The Conflict and Environment Observatory
[3]University of Leeds

March 04, 2024

Abstract

Thousands of people are injured every year from explosive remnants of war which include unexploded ordnance (UXO) and abandoned ordnance. UXO has negative long-term impacts on livelihoods and ecosystems in contaminated areas. Exact locations of remaining UXO are often unknown as survey and clearance activities can be dangerous, expensive and time-consuming. In Vietnam, Lao PDR and Cambodia, about 20% of the land remains contaminated by UXO from the Vietnam War. Recently declassified historical KH-9 satellite imagery, taken during and immediately after the Vietnam War, now provides an opportunity to map this remaining contamination. KH-9 imagery was acquired and orthorectified for two study areas in Southeast Asia. Bomb craters were manually labeled in a subset of the imagery to train convolutional neural networks (CNNs) for automated crater detection. The CNNs achieved a F1-Score of 0.61 and identified more than 500,000 bomb craters across the two study areas. The detected craters provided more precise information on the impact locations of bombs than target locations available from declassified U.S. bombing records. This could allow for a more precise localization of suspected hazardous areas during non-technical surveys as well as a more fine-grained determination of residual risk of UXO. The method is directly transferable to other areas in Southeast Asia and is cost-effective due to the low cost of the KH-9 imagery and the use of open-source software. The results also show the potential of integrating crater detection into data-driven decision making in mine action across more recent conflicts.

Hosted file

`manuscript.docx` available at https://authorea.com/users/750625/articles/721703-detecting-vietnam-war-bomb-craters-in-declassified-historical-kh-9-satellite-imagery

Detecting Vietnam War Bomb Craters in Declassified Historical KH-9 Satellite Imagery

Philipp Barthelme[1], Eoghan Darbyshire[2], Dominick Spracklen[3], and Gary R Watmough[1,4,5]

[1] School of Geosciences, University of Edinburgh, Edinburgh, United Kingdom.

[2] Conflict and Environment Observatory, Mytholmroyd, United Kingdom.

[3] School of Earth and Environment, University of Leeds, Leeds, United Kingdom.

[4] Global Academy of Agriculture and Food Systems, University of Edinburgh, United Kingdom

[5] Novel Data Ecosystems for Sustainability (NODES) Research Group. Advancing Systems Analysis Program, International Institute for Applied Systems Analysis (IIASA), Austria

Corresponding author: Philipp Barthelme (philipp.barthelme@ed.ac.uk)

Key Points:

- We detect bomb craters in declassified Vietnam War-era satellite imagery to understand remaining unexploded ordnance contamination
- We find that crater appearance differs by location and changes over time which affects the detection accuracy
- We show that detected craters are more precise than bombing records which can improve data-driven decision making in mine action

Abstract

Thousands of people are injured every year from explosive remnants of war which include unexploded ordnance (UXO) and abandoned ordnance. UXO has negative long-term impacts on livelihoods and ecosystems in contaminated areas. Exact locations of remaining UXO are often unknown as survey and clearance activities can be dangerous, expensive and time-consuming. In Vietnam, Lao PDR and Cambodia, about 20% of the land remains contaminated by UXO from the Vietnam War. Recently declassified historical KH-9 satellite imagery, taken during and immediately after the Vietnam War, now provides an opportunity to map this remaining contamination. KH-9 imagery was acquired and orthorectified for two study areas in Southeast Asia. Bomb craters were manually labeled in a subset of the imagery to train convolutional neural networks (CNNs) for automated crater detection. The CNNs achieved a F1-Score of 0.61 and identified more than 500,000 bomb craters across the two study areas. The detected craters provided more precise information on the impact locations of bombs than target locations available from declassified U.S. bombing records. This could allow for a more precise localization of suspected hazardous areas during non-technical surveys as well as a more fine-grained determination of residual risk of UXO. The method is directly transferable to other areas in Southeast Asia and is cost-effective due to the low cost of the KH-9 imagery and the use of open-source software. The results also show the potential of integrating crater detection into data-driven decision making in mine action across more recent conflicts.

Plain Language Summary

Every year, thousands of people are injured or killed by unexploded weapons from previous wars. In Vietnam, Lao PDR and Cambodia, unexploded bombs from the Vietnam War remain in about 20% of the land. Clearing this area is expensive and could take decades, requiring prioritization and risk management. To identify the most affected areas, we used machine learning methods to find bomb craters in satellite images within two study areas. As bomb craters often change appearance or completely disappear over time, making them difficult to detect in today's satellite images, we used recently declassified U.S. satellite images, taken during the Vietnam War. We found that the detected crater locations are more precise than target locations from bombing records, as they show where the bombs actually exploded, which we found can be kilometers away from their recorded targets. Although the presence of bomb craters means that the corresponding bombs exploded, any unexploded bombs from the same bomb strike are likely to be located nearby. Detected crater locations can therefore be used to more precisely define the areas where unexploded bombs are suspected to remain, which can help to make subsequent clearance activities more efficient and risk management more reliable.

1 Introduction

Unexploded ordnance (UXO) refers to explosive munitions, including bombs, artillery projectiles and cluster submunitions that have been deployed during military conflicts but did not explode. UXO continues to present significant humanitarian and environmental challenges. In 2022 alone, the United Nations Mine Action Service (UNMAS) reported more than 3,000 casualties from explosive remnants of war, which include UXO and abandoned explosive ordnance, across 15 countries and numbers of UXO are increasing due to ongoing conflicts such as in Ukraine (Cluster Munition Coalition, 2023; UNMAS, 2019). UXO has negative long-term impacts on public health, livelihoods and ecosystems in contaminated areas (Frost et al., 2017; Hofmann & Juergensen, 2017; E. Lin et al., 2020; Nguyen, 2020; Ounmany & Andriesse, 2018). Moreover, the removal of UXO remains technically challenging, expensive and hazardous, particularly in conflict and post-conflict environments where access to reliable data on contamination is limited.

Mainland Southeast Asia has one of the highest UXO contamination rates in the world, mainly originating from the aerial bombardment by the U.S. military during the Vietnam War, also known as the American War in Vietnam or the Second Indochina War which took place between 1955 and 1975 (Martin et al., 2019). During the war, the U.S. Air Force dropped approximately eight million tons of bombs on the countries of Vietnam, Cambodia and Lao PDR (Anderson, 2002; High et al., 2013). Today, about 20% of the land in these countries is thought to still be contaminated by UXO (Martin et al., 2019). However, the exact locations and extents of contaminated areas mostly remain unknown, despite being essential for an efficient allocation of limited resources for UXO clearance. Non-technical survey is commonly used as a first step to identify contaminated land and categorize it into suspected or confirmed hazardous areas. This approach relies on the collection and analysis of all available data about possible explosive ordnance contamination in an area, including historical records such as locations of army bases, battle areas and bombing targets. As it is cheaper than technical survey and clearance, which rely on expensive technical assets to be deployed to the field, an accurate non-technical survey can ensure the most efficient allocation of limited resources (Bold & Avenell, 2021; E. Lin et al., 2020; UNMAS, 2019).

U.S. bombing records are one of the most comprehensive data sources used for non-technical survey in Southeast Asia. In 2016, the United States Department of Defense released these records to the public as part of the Theater History of Operations (THOR) data, an attempt to record all air operations by the United States since World War I. The THOR data includes the geographical coordinates of target locations, the type and number of weapons dropped on each target and the time of the attack. The bombing records have been a valuable data source for non-technical surveys (Bold & Avenell, 2021) and for research into the political, economic and health impacts of the Vietnam War (D. T. Le et al., 2022; K. Le & Nguyen, 2020; Yamada & Yamada, 2021). However, High et al. (2013) suggest the bombing data should only be used as one source among many, after identifying multiple issues, including missing, corrupted and actively falsified records. An overview of THOR bombing targets in Southeast Asia during the Vietnam War is shown in Figure 1a.

Remote sensing data can provide a valuable alternative data source where bombing records are unavailable or inaccurate (Bennett et al., 2022). Lin et al. (2020) used recent, very high resolution (< 1 m) satellite imagery to detect bomb craters from the Vietnam War in Cambodian agricultural land. However, detecting bomb craters from past conflicts in more recent satellite images can be challenging as the appearance of bomb craters changes over time due to erosion, vegetation growth and human intervention (E. Lin et al., 2020). Historical aerial wartime imagery has been used as an alternative to detect and analyze World War II bomb craters in Europe (Clermont et al., 2019; Kruse et al., 2019; Waga et al., 2022), but its availability is often restricted to small areas. Declassified historical U.S. satellite imagery (USGS EROS Center, 2018), taken during and immediately after the Vietnam War, now presents an opportunity to overcome some of these challenges. The KH-4a/b CORONA missions provide high resolution imagery (1.8-2.8 m) between 1963 and 1972 which, since its declassification in 1995, has been used in a variety of applications that range from the discovery of archaeological sites to land cover change detection (Deshpande et al., 2021; Lasaponara et al., 2018; Nita et al., 2018). Recently, it was used to classify land affected by bombing in a part of Quang Tri province, Vietnam (Munteanu et al., 2024). The KH-9 HEXAGON stereo-panoramic imagery provides almost complete coverage of the Earth's land area between 1971 and 1984 at a spatial resolution of 0.6-1.2 m. Due to its recent declassification in 2011 and the technical challenges associated with orthorectifying the imagery (Zhou et al., 2021), researchers have only recently begun to explore its use in a diverse range of applications such as archaeology (Hammer et al., 2022) and glaciology (Ghuffar et al., 2023).

Previous studies on the automatic detection and counting of bomb craters in remotely sensed imagery have relied on methods developed for detecting extra-terrestrial craters on planetary surfaces (Clermont et al., 2019; E. Lin et al., 2020). In this field, convolutional neural networks (CNNs) are increasingly replacing applications that rely on the extraction of manually specified features such as crater shape and shadows. U-Nets, a type of CNN architecture originally developed for segmenting medical imagery (Ronneberger et al., 2015), have been successfully applied to segment extra-terrestrial craters (Chen et al., 2023; Silburt et al., 2019) and more recently to detect artillery craters in Ukraine (Duncan et al., 2023). To achieve instance segmentation, a method for identifying individual instances of an object, methods are often adjusted by introducing a boundary class (Duncan et al., 2023) or by using template matching on the semantic segmentation product (Chen et al., 2023; Silburt et al., 2019).

Our study was structured in the following way. First, we acquired KH-9 imagery for two study areas in Southeast Asia and orthorectified the imagery using open-source tools. We manually labeled craters on a subset of the imagery and categorized them based on their appearance. To automate crater detection, we used an instance segmentation workflow using CNNs with a U-Net architecture. We then analyzed how the model performance varies for different crater appearances. Finally, we compared detected crater locations to U.S. bombing records, identifying multiple issues with the bombing records in the process. Our results show that craters visible in the KH-9 imagery provide more precise information about where bombs landed than currently used bombing records. Additionally, our findings demonstrate how methods to automatically detect these craters can improve data-driven decision making within the mine action sector in Southeast Asia.

2 Materials and Methods

2.1 Study areas

Two study areas across Southeast Asia were selected (Figure 1). The first study area covers a total of 4,148 km^2 of Quang Tri (QT) province, the most heavily bombed province in Vietnam during the war (Miguel & Roland, 2011), as it contained the 17^{th} parallel, the dividing line between North and South Vietnam at the time. The second study area, here referred to as the tri-border area (TBA), is located around the meeting point of the borders of Vietnam, Lao PDR, and Cambodia. Encompassing 17,285 km^2 of predominantly mountainous and densely vegetated land, the TBA contained sections of the Ho Chi Minh Trail, the principal supply route for the North Vietnamese Army, including a vital entry point of the trail into South Vietnam in Kon Tum province. The KH-9 images were taken on November 4, 1972 (TBA) and on March 20, 1973 (QT province).

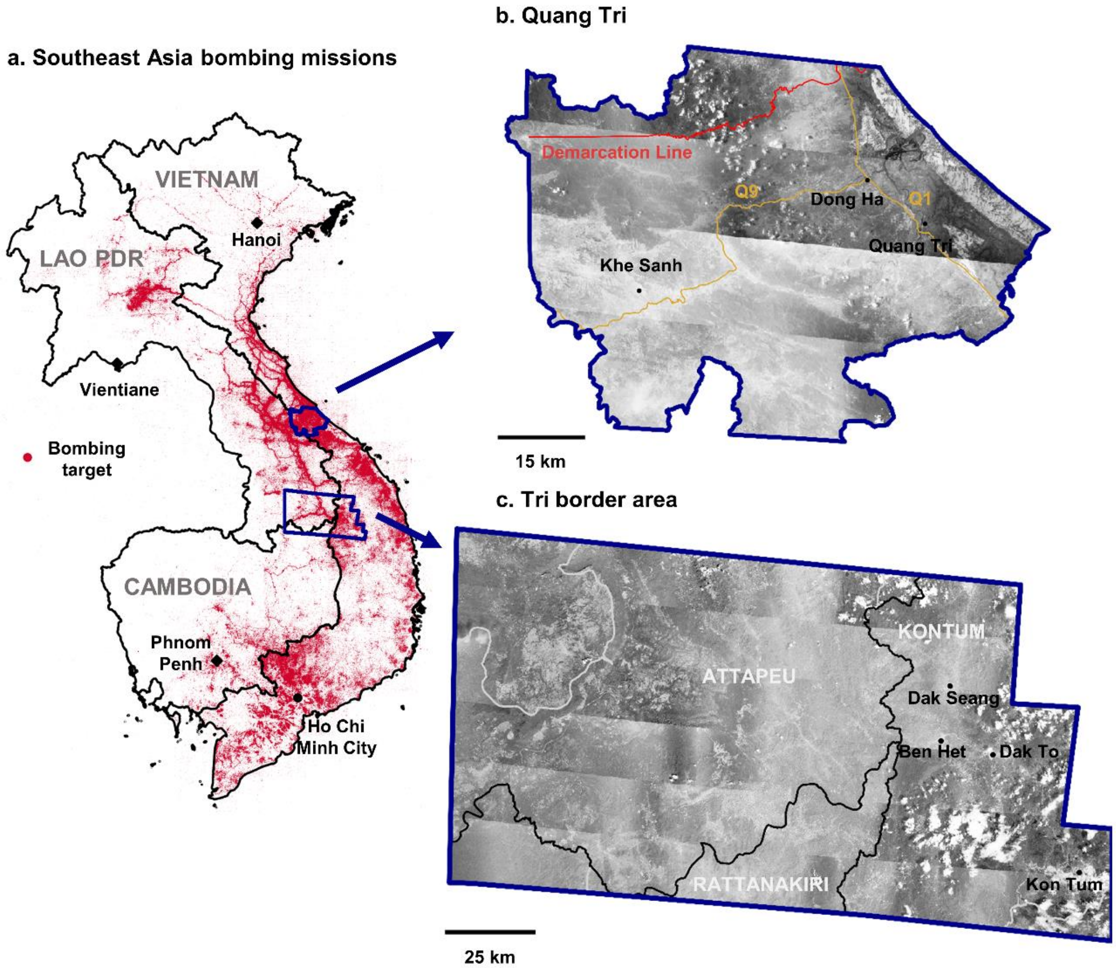


Figure 1. (a) THOR bombing targets over Southeast Asia during the Vietnam War. (b) and (c) show the two study areas, including points of interest during the war. The cities of Dong Ha, Quang Tri, Dak To and Kon Tum were locations of larger military bases whereas Khe Sanh, Dak Seang and Ben Het contained smaller military camps.

2.2 Processing the KH-9 imagery

A total of 20 KH-9 images, forming 10 stereo pairs of forward and aft looking cameras, were used for the study. The U.S. Geological Survey provided photogrammetric film scans of the archived KH-9 film sources at a resolution of 7 microns and a cost of 30$ per image. Previously digitized images, now including all images used in this study, are available at no cost via the Earth Explorer platform.

The film scans were provided in multiple sections and were not georeferenced. The open-source Nasa Ames Stereo Pipeline (ASP) (Beyer et al., 2021) was used to process and orthorectify the imagery. The ASP implements a rigorous camera model including motion compensation (Sohn et al., 2004) for the panoramic cameras used by the KH-9 satellites. We adapted the example workflow described in section 8.26 of the ASP manual (Beyer et al., 2021), as we integrated manual ground control points (GCPs) to improve accuracy.

First, image parts were stitched together and cropped to the image extent using the *image_mosaic* and *historical_helper* tools in ASP. QGIS (QGIS Association, 2023) and Google Earth imagery were used to identify approximately 15 ground control points (GCPs) per image. The GCPs were used to initialize intrinsic and extrinsic camera parameters which were further optimized using a joint bundle adjustment for each stereo pair. The optimized camera parameters were used to project each image onto a digital elevation model (NASA Shuttle Radar Topography Mission (SRTM), 2013) at a resolution of one meter per pixel using the *mapproject* tool in ASP.

The resulting images were cropped to the study areas introduced in Section 2.1. The QT imagery was mosaicked into one image, while the TBA images, being larger in size, were not mosaicked. A further 60 validation GCPs (QT province: 20, TBA: 40) were collected for validation of the orthorectification process and showed a mean absolute horizontal error of 7.0 m (25^{th} percentile: 3.4 m, median: 5.8 m, 75^{th} percentile: 8.8 m) for QT province and 17.5 m (25^{th} percentile: 8.6 m, median: 13.7 m, 75^{th} percentile: 21.4 m) for the TBA.

2.3. Labeling of bomb craters

The processed KH-9 imagery was divided into image tiles with a width and height of 256 pixels. From the QT imagery, 1,000 random image tiles were chosen and divided into sets of 600 for training, 200 for validation, and 200 for testing. From the TBA imagery, 1,400 tiles were selected, with 600 allocated for training, 200 for validation, and 600 for testing. The decision to increase the number of test tiles for the TBA was driven by its lower density of bomb craters, aiming to ensure a more representative test score.

Craters visible in the selected image tiles were manually labeled if they were larger than 25 pixels (equivalent to 25 m^2). Smaller ground features were excluded as they were difficult to reliably identify given the image resolution and quality. The threshold of 25 pixels was selected based on visual inspection. Each labeled crater was assigned one of five classes based on its appearance in the imagery which varied substantially (Figure 2).

Labeling proved particularly challenging in mountainous areas with heavy vegetation and in areas featuring houses, trees or graves that could resemble craters in the imagery. Where necessary, the context visible in the KH-9 and current satellite imagery was used to make a better-informed decision. Notably, the crater prevalence was much lower for the TBA where 964 craters were identified compared to 10,132 craters in QT province. Additional details on the crater labeling and the different crater classes are provided in the Supplementary Materials.

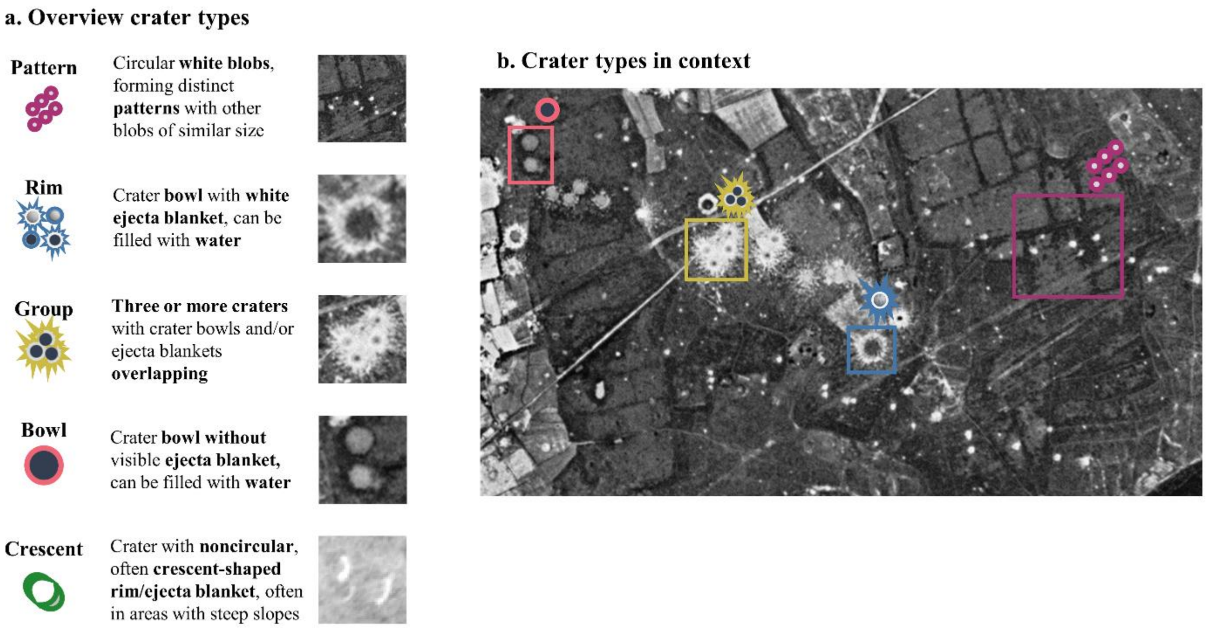


Figure 2. (a) Different crater types defined based on their appearance characteristics in the KH-9 imagery. (b) Examples of the crater types in context for an area in Quang Tri province.

2.4. Detection of bomb craters

An instance segmentation workflow was used to predict individual craters in the imagery. The instance segmentation was implemented as a semantic segmentation problem by adding a boundary class, an approach commonly used in biomedical applications such as nucleus segmentation (Caicedo et al., 2019), where large amounts of densely packed objects have to be separated. For this approach, the area of each labeled crater was expanded by two pixels which were assigned to the new boundary class. All crater pixels that were touching neighboring craters were also labeled as boundary pixels.

2.4.1. Neural network architecture and training

A U-Net with a Resnet50 backbone, pre-trained on the Imagenet dataset, was used for the segmentation (Deng et al., 2009; He et al., 2015; Ronneberger et al., 2015). While multiple improvements to the standard U-Net architecture have been suggested, in most settings they only lead to minor or no accuracy improvements at a larger computational cost (Gut et al., 2022; Kugelman et al., 2022; Wang & Miao, 2022). Therefore, instead of comparing different model architectures, the analysis in this paper focuses on different bomb crater appearances, which have a large impact on model accuracy, and the comparison of detected craters with historical bombing records.

The model was implemented using the *pytorch* and *segmentation_models_pytorch* packages in Python (Iakubovskii, 2019; Paszke et al., 2019). An initial model was trained using data from both study areas before the model was fine-tuned for each study area independently, using only the training data for the respective study area. Min-max scaling was applied to individual image tiles. During model training, the images were augmented by applying random vertical and horizontal flips as well as random brightness and contrast adjustments. As the pre-trained model expected color images with three channels as input, whereas the KH-9 images are grayscale with a single channel, an additional layer was added in front of the pre-trained model to map from one to three channels.

The models were trained on a Nvidia RTX 2060 GPU using an Adam optimizer (Kingma & Ba, 2014) and a focal loss function (T.-Y. Lin et al., 2017), which assigned more weight to training examples that were not well classified. Focal loss has been shown to work well for imbalanced data (T.-Y. Lin et al., 2017; Mulyanto et al., 2021) which was a problem here as more than 99% of all labeled pixels were background pixels. A focal loss alpha value of 1 was used for the background class and 3 for all other classes based on the model performance on the validation images. A batch size of 8 and a learning rate of 1e-3 was used during initial model training and the learning rate was reduced to 1e-5 for the fine-tuning of each study area. Early stopping was used to stop model training if the validation loss did not decrease for 50 epochs in a row.

2.4.2. Semantic segmentation evaluation

The segmentation results were evaluated using a pixel-to-pixel comparison on the test images. We used precision, recall and F1-score which are commonly applied in settings of class imbalance and which are defined as:

$$Precision = \frac{TP}{TP + FP} \#(1)$$

$$Recall = \frac{TP}{TP + FN} \#(2)$$

$$F_1 = \frac{2TP}{2TP + FN + FP} , \#(3)$$

where *TP* denotes true positives, *FN* denotes false negatives and *FP* denotes false positives. We calculated these metrics for each individual crater class, and additionally calculated one combined score that only considers whether a pixel had been correctly identified as a crater pixel, even if the crater class of predicted and labeled pixels differed.

2.4.3. Instance segmentation

Multiple post-processing steps are applied to transform the semantic segmentation output into individual crater instances (Figure 3). Connected crater pixels were considered as one crater instance even if they belonged to different crater classes. Pixels of class *Boundary* were treated as background pixels at this stage. Each predicted crater instance was assigned the majority class of its pixels. All predicted crater instances smaller than 25 pixels were removed.

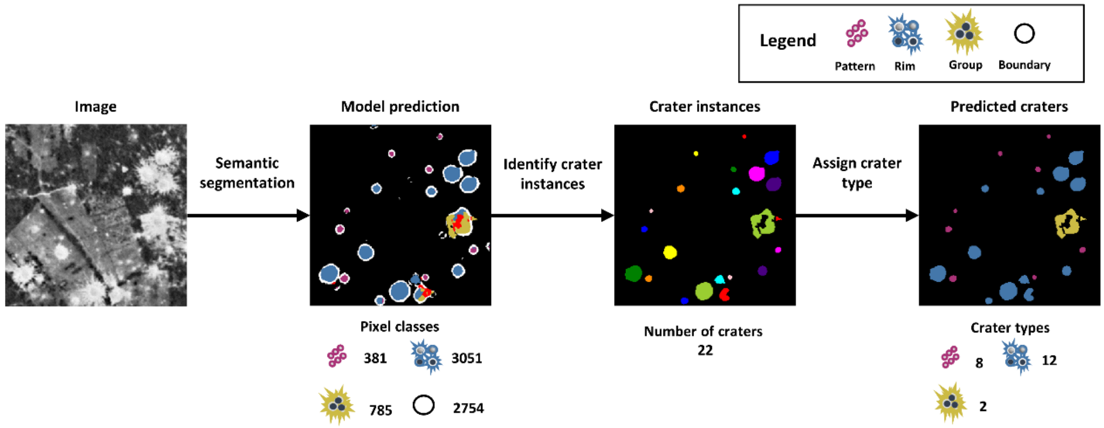


Figure 3. Crater prediction and post-processing workflow.

The accuracy of crater instances was evaluated using the metrics described in Section 2.4.2. A predicted crater (A) was considered correct if it had an *Intersection over Union* (*IoU*) of 0.5 or more with a labeled crater (B), where *IoU* is defined as:

$$IoU(A, B) = \frac{A \cap B}{A \cup B}\ . \#(4)$$

Accuracy scores were calculated for each individual crater type and for a combined crater class that only considered whether a crater instance had been correctly identified even if the crater class of the predicted and labeled craters differed.

2.4.4. Model prediction

The trained models were applied to the entire study areas using a sliding window approach with an overlap of 64 pixels. Only the center 192×192 pixels of each predicted 256×256 image tile were retained to avoid artefacts and improve performance at tile edges. When identifying individual crater instances on the predicted segmentation masks the tile size of 1024×1024 with an overlap of 512 pixels was used to avoid mistakenly separating large craters that crossed one or more image tiles. As the images, and therefore crater predictions, in the second study area were overlapping, we only kept predicted craters of one of the images in these overlapping areas. This was not necessary for QT province, where the KH-9 images were mosaicked before crater detection, resulting in the same outcome.

2.5. THOR bombing data

We identified multiple issues with the THOR bombing data, including (1) coordinates being labeled as using the WGS84 datum while our analysis suggested they were provided in the Indian 1960 datum; (2) double counting of B-52 bombing missions from 1971 onwards; (3) wrongly assigned mission functions resulting in wrongly assigned kinetic and non-kinetic mission classifications; and (4) missing weapon types for a large proportion of the records. We corrected the THOR records to the best of our knowledge to make them usable for the purpose of our research, which compared the bombing on a large scale, but note that some limitations and uncertainty remain. Details of the identified issues and applied corrections, including checks to the robustness of our results based on alternative processing, are provided in the Supplementary Material.

The analysis was limited to large aircraft bombs which would result in craters larger than 25 m^2. All records of bombing that occurred after the respective KH-9 images were taken were dropped. The resulting records are referred to as *total bombing*. The data were further split into (1) bombs dropped within the year before the respective KH-9 images were taken (*previous year bombing*) and (2) bombs dropped more than a year before the imagery was taken (*bombing before previous year*). The resulting numbers of bombs dropped were directly compared to the number of detected craters in each study area. Additionally, aggregated counts of detected craters and bombs dropped for grid cells of various cell sizes between 100 m and 4 km were compared using the Spearman correlation coefficient r (Schober et al., 2018). To allow for a direct comparison between the number of detected craters and the number of bombs dropped during *previous year bombing*, excluding the influence of older craters, a distinct analysis was undertaken. This analysis focused on grid cells (2 km × 2 km) within QT province, where *previous year bombing* constituted at least 90% of *total bombing*.

3 Results and Discussion

3.1. Model evaluation

The trained models achieved an F1-Score of 0.61 (precision: 0.67, recall: 0.56) when predicting craters of all types across the test sets and predicted a total of 541,398 craters (QT: 442,157, TBA: 99,241) across the full study areas (Figure 4). The model performance differed between the two study areas with an F1-Score of 0.64 for QT province and 0.44 for the TBA (Table 1). We present detailed metrics by study area and crater types in Table 1. Most of the predicted craters were of type *Pattern* (QT: 229,467, TBA: 67,985), *Rim* (QT: 91,112, TBA: 9,995) and *Crescent* (QT: 71,364, TBA: 16,836). The model only predicted a small number of craters of type *Group* (QT: 9,645, TBA: 46) and *Bowl* (QT: 40,569, TBA: 4,379). Figure 5 shows the detected crater locations by crater type for the QT study area.

Table 1. Bomb crater detection results showing F1-score (precision/recall) and the number of labeled craters N in the test data. For the Craters category, all crater classes are considered as one combined crater class.

	Craters	Pattern	Rim	Group	Crescent	Bowl	Boundary	Background
Quang Tri								
Pixels	**0.65 (0.70/0.61)**	0.63 (0.62/0.64)	0.59 (0.60/0.58)	0.13 (0.45/0.07)	0.35 (0.31/0.40)	0.27 (0.35/0.22)	0.45 (0.44/0.46)	0.99 (0.99/0.99)
Craters (IOU > 0.5)	**0.64 (0.68/0.60) N=1712**	0.70 (0.68/0.73) N=748	0.55 (0.58/0.52) N=449	0.07 (0.33/0.04) N=247	0.30 (0.25/0.37) N=111	0.24 (0.30/0.20) N=157	-	-
Tri-border area								
Pixels	**0.41 (0.61/0.31)**	0.53 (0.56/0.50)	0.38 (0.67/0.27)	0.00 (0.00/0.00)	0.17 (0.21/0.14)	0.06 (0.47/0.03)	0.29 (0.39/0.23)	1.00 (1.00/1.00)
Craters (IOU > 0.5)	**0.44 (0.59/0.35) N=314**	0.58 (0.59/0.57) N=142	0.37 (0.64/0.26) N=54	0.00 (0.00/0.00) N=15	0.20 (0.25/0.17) N=36	0.03 (0.20/0.02) N=67	-	-

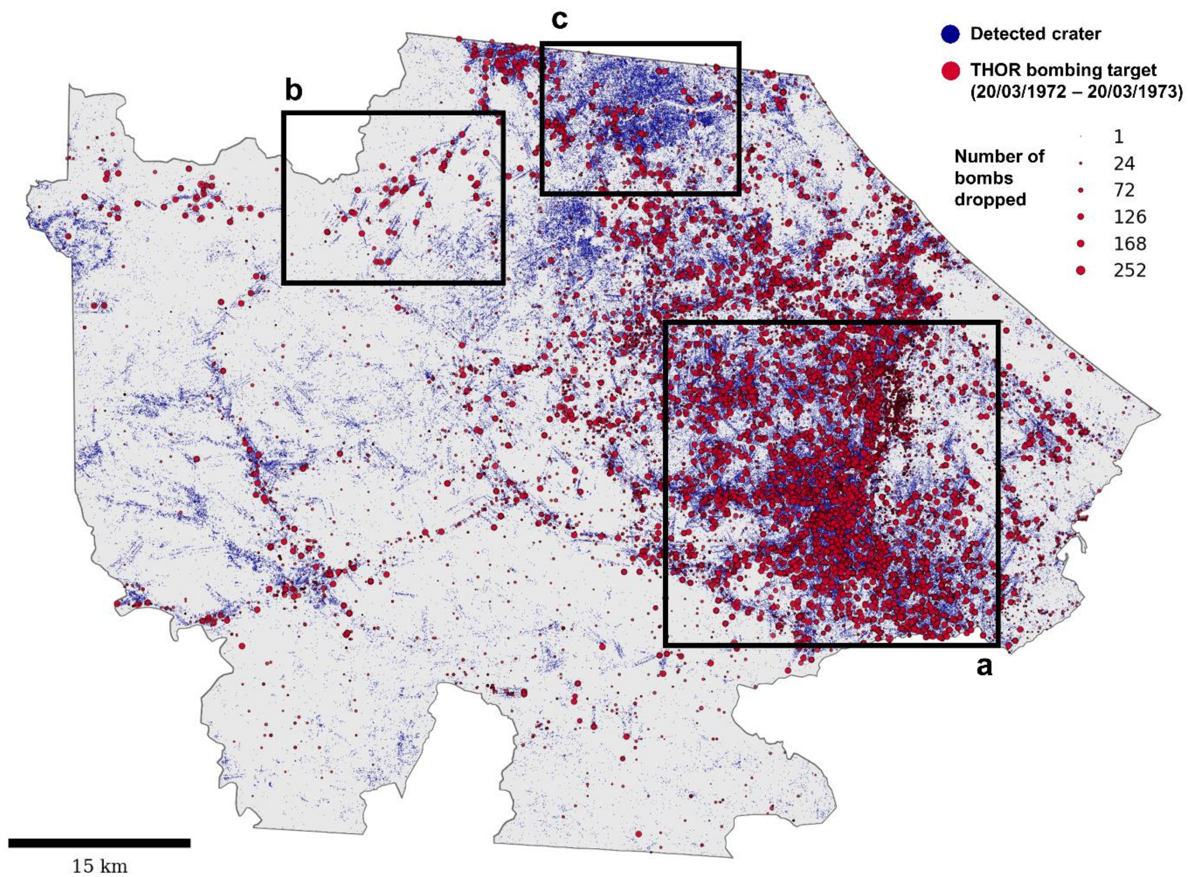


Figure 4. Comparison of predicted bomb craters (blue) and THOR bombing targets (red) during the year preceding the KH-9 image acquisition in Quang Tri province. (a) shows a high density of bomb craters and bombs dropped close to Quang Tri city. (b) shows multiple lines of bomb craters matching B-52 bombing targets recorded in THOR. (c) shows an area with large amounts of craters but little bombing during the year before the KH-9 images were taken, indicating the craters originated from earlier in the war.

3.1.1. Model performance across study areas

The difference in F1-Scores for the two study areas is likely partly due to the lower prevalence of craters in the TBA where only 1 in 1434 pixels were crater pixels compared to 1 in 90 for QT province. The lower prevalence results in a smaller number of labeled craters and a higher influence of every false positive crater prediction on the evaluation metrics, which has been identified as a challenge in previous research on bomb crater detection (Clermont et al., 2019; Lin et al., 2020). As a random sample of images was used in each study area, the test data had a realistic class distribution, and the results for the TBA reflect the difficulty of predicting bomb craters over large, mostly unaffected areas. Further, land cover can influence the accuracy of predictions; in the TBA, land cover mostly consisted of heavily vegetated land and mountainous terrain, with only small amounts of agricultural land in which bomb craters are generally easier to identify and segment (Duncan et al., 2023).

3.1.2. Model performance across crater types

Model performance varied between the different crater types (Table 1), with higher F1-scores for craters of type *Pattern* and *Rim* compared with *Group*, *Crescent* and *Bowl*. This could be attributed to the lower prevalence for these crater types, resulting in fewer training data for the model to learn from especially in the TBA imagery, which is why we focus the rest of the discussion of crater types on the results for QT province.

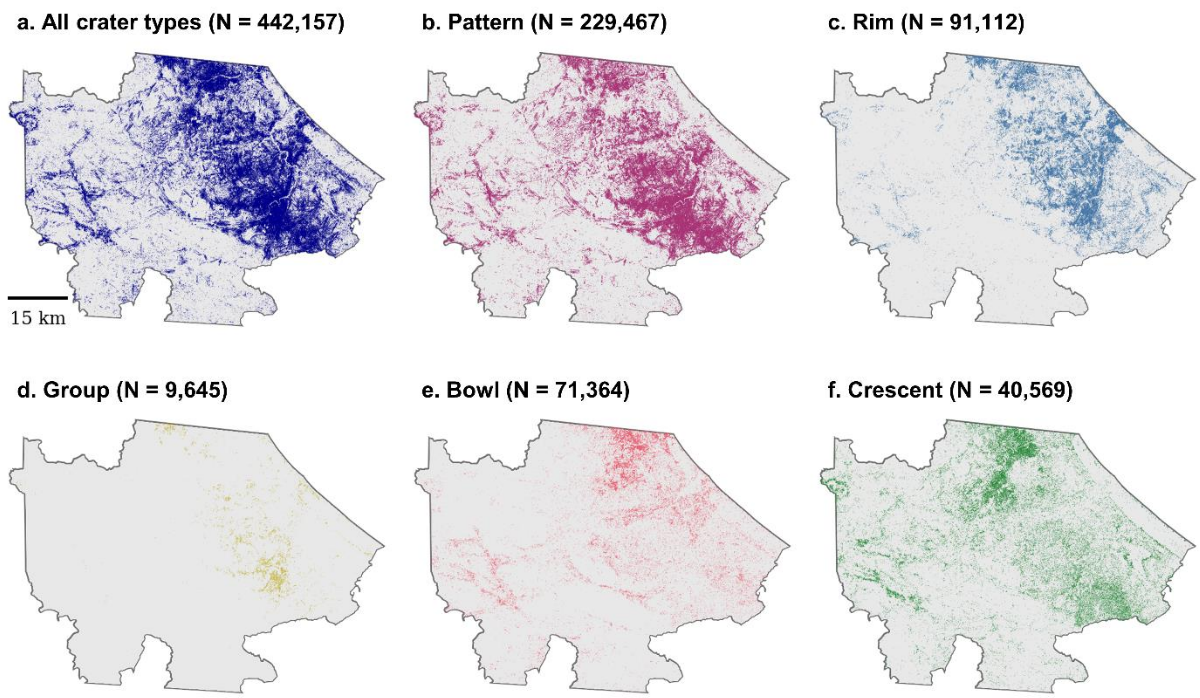


Figure 5. Predicted craters by crater class in Quang Tri province. The total number of detected craters N is provided for each class. There is a clear difference in the distribution of the crater classes. Rim and Bowl craters were mostly located in the paddy fields closer to the coast, where the Rim craters seem to match better with previous year bombing. Group craters were rare and only predicted in very specific locations that have seen the heaviest bombing. Pattern and Crescent craters were spread across the whole study area.

Visual inspection and the pixel level accuracy assessment highlighted that the model detected *Group* craters in the correct areas (Figure 5). However, the model did not accurately separate individual crater instances, a challenge that we also encountered during crater labeling. One way to address this could be to use an area-based approach that treats overlapping craters as one object and uses the total covered area instead of the crater count as a metric. *Crescent* craters were often located in areas with steep slopes and dense vegetation which meant that the appearance of these craters varied substantially, making reliable labeling difficult. *Bowl* craters were often old craters that had eroded and blended into the surroundings, which made labeling and detection challenging. These craters often occurred along rivers and canals where they were filled with water and only visible as dark circular blobs that could be confused with other ground features like trees. Therefore, *Crescent* or *Bowl* craters would be easier to detect in images taken closer to the date of the bombing.

3.2. Comparison of detected bomb craters with THOR bombing data

The THOR bombing records show that around 1 million bombs (QT: 654,730, TBA: 321,504) were dropped across the two study areas during the year preceding the KH-9 image acquisition (*previous year bombing)* and more than 3 million bombs (QT: 2.23 million, TBA: 1.13 million) during the entire conflict before the respective KH-9 images were taken (*total bombing*). Comparisons between detected craters and number of dropped bombs over grid cells of 2 km × 2 km (Table 2) indicated that craters were positively correlated with *previous year bombing* (QT: r =0.76, TBA: r =0.51) and *total bombing* (QT: r =0.58, TBA: r =0.51) and correlation coefficients increased with grid cell size (Figure 6).

A visual comparison showed that detected craters were located close to THOR target locations and were often organized in lines of craters characteristic for the B-52 bombing strikes (Figure 4). The predicted crater locations are overlayed with aggregated bombing data for QT province (grid size: 2 km × 2) and the TBA (grid size: 4 km × 4 km) in Figure 7. In grid cells in QT province for which more than 90% of *total bombing* happened during the year before KH-9 image acquisition, the model detected a total of 157,846 craters, accounting for 46% of the 344,135 bombs dropped during *previous year bombing* (44% of *total bombing*).

Table 2. Spearman correlation coefficients between detected craters and number of bombs dropped (THOR) aggregated across grid cells of 2 km × 2km. For the Craters category detected craters of all crater classes were aggregated.

	Craters	Pattern	Rim	Group	Crescent	Bowl	Number of bombs dropped
Quang Tri							
Previous year bombing	**0.76**	0.74	0.78	0.68	0.52	0.62	654,730
Total bombing	**0.58**	0.55	0.46	0.33	0.61	0.41	2,232,280
Number of detected craters	**442,157**	229,467	91,112	9,645	71,364	40,569	-
Tri-border area							
Previous year bombing	**0.51**	0.52	0.42	0.11	0.47	0.34	321,504
Total bombing	**0.51**	0.51	0.40	0.09	0.44	0.34	1,133,025
Number of detected craters	**99,241**	67,985	9,995	46	16,836	4,379	-

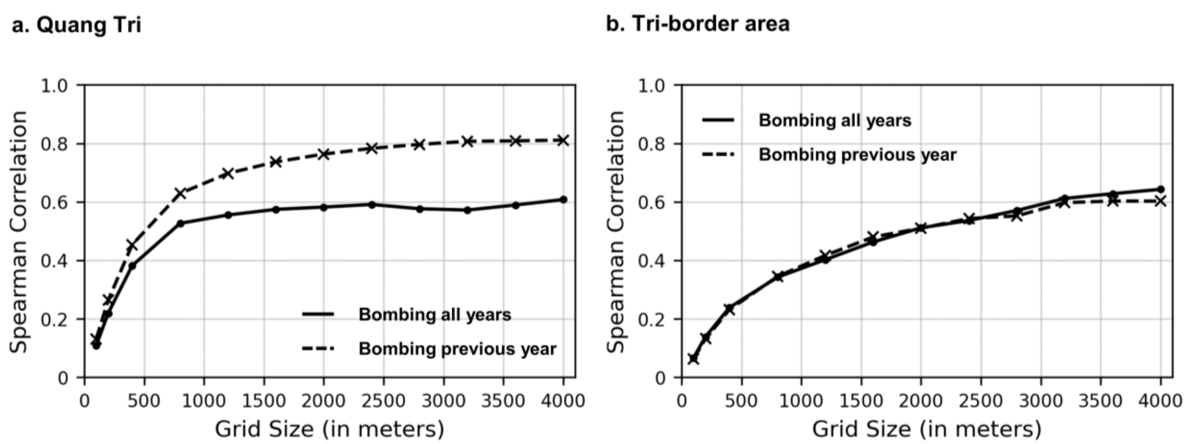


Figure 6. Spearman correlations of the aggregated number of detected craters against bombs dropped (THOR) within grid cells of multiple sizes.

3.2.1. Spatial precision

Craters identified in the KH-9 imagery can offer more precise information about potential UXO locations compared to the THOR data. While each THOR record is confined to a single target location, it can encompass tens or hundreds of dropped bombs. Figure 8, depicting three target locations of B-52 bombing missions, shows resulting craters spanning several kilometers. Only few of these craters lie within a hundred-meter radius of the target location which explains the low correlations between detected craters and dropped bombs for smaller grid sizes (Figure 6). In the instance illustrated in Figure 8, our estimation indicates that identifying the impact crater locations from the B-52 bomb strikes reduces the potential area for locating unexploded bombs from those strikes to about 9% of the area derived from the THOR target locations alone, as unexploded bombs are likely to be located near the lines of craters.

Moreover, the KH-9 imagery can be useful to identify and correct errors in the THOR data. The imagery in Figure 8 revealed a discrepancy with the THOR data, where no nearby craters were visible for one target location. According to the THOR records, this mission had been diverted with some bombs supposedly dropped on the target visible in Figure 8 and the remainder on a second target. However, the KH-9 imagery suggests it is more likely that all bombs were dropped at the second target and none at the first. This highlights the advantages of having multiple independent data sources that can be cross-referenced for a more thorough analysis.

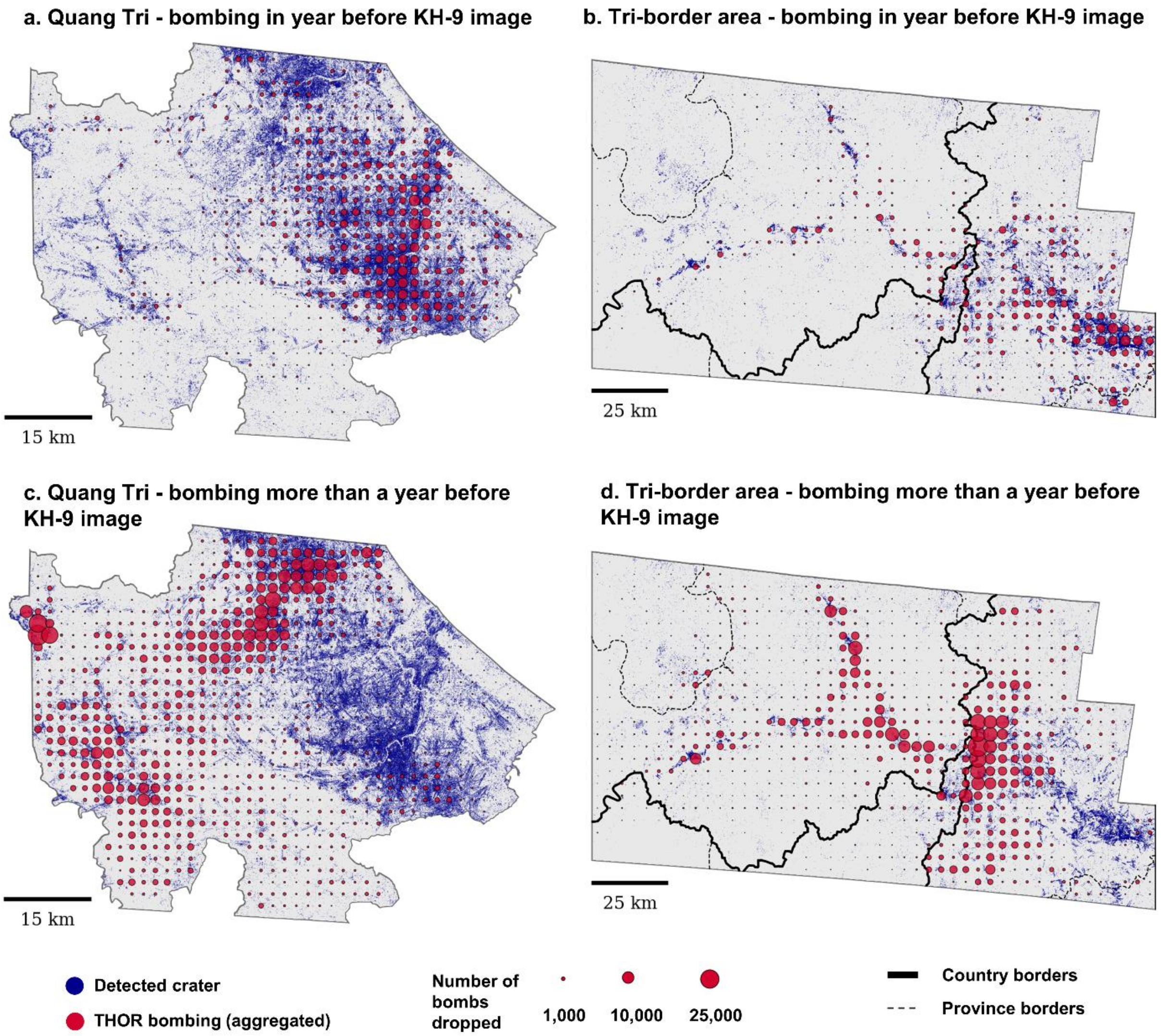


Figure 7. Comparison of predicted bomb craters (blue) and THOR bombs dropped (red) aggregated across grid cells of 2 km × 2 km for Quang Tri province and 4 km × 4 km for the tri-border area. Bombs dropped during the year preceding the KH-9 image acquisition are shown in (a) and (b) whereas (c) and (d) show all bombing that happened more than a year before the image acquisition.

3.2.2. Temporal analysis

Bomb craters can become increasingly difficult to detect from space over time. In Southeast Asia, with its dense rainforests and regular flooding, craters can quickly become covered up by vegetation, deformed by erosion or filled up by humans (E. Lin et al., 2020). Our analysis underscores that these effects impact detection results even after short periods of time, not only limiting the utility of current satellite imagery but emphasizing the need for additional imagery taken during the earlier stages of the war.

Our models detected a high concentration of craters near the cities of Quang Tri and Kon Tum, which were subject to heavy bombing during the year preceding the KH-9 image acquisition. Comparatively fewer craters were detected in areas targeted earlier during the war (Figure 7). This is reflected by a higher correlation of detected craters with *previous year bombing* compared to *total bombing* in QT province, albeit not for the TBA (Table 2). Notably, we encountered challenges labeling and detecting partly eroded craters formed by bombs dropped earlier in the war, which we associated with the crater types *Bowl* and *Crescent*. To mitigate this bias towards areas bombed later in the war, we propose the use of the CORONA imagery captured during the earlier phases of the conflict (Munteanu et al., 2024).

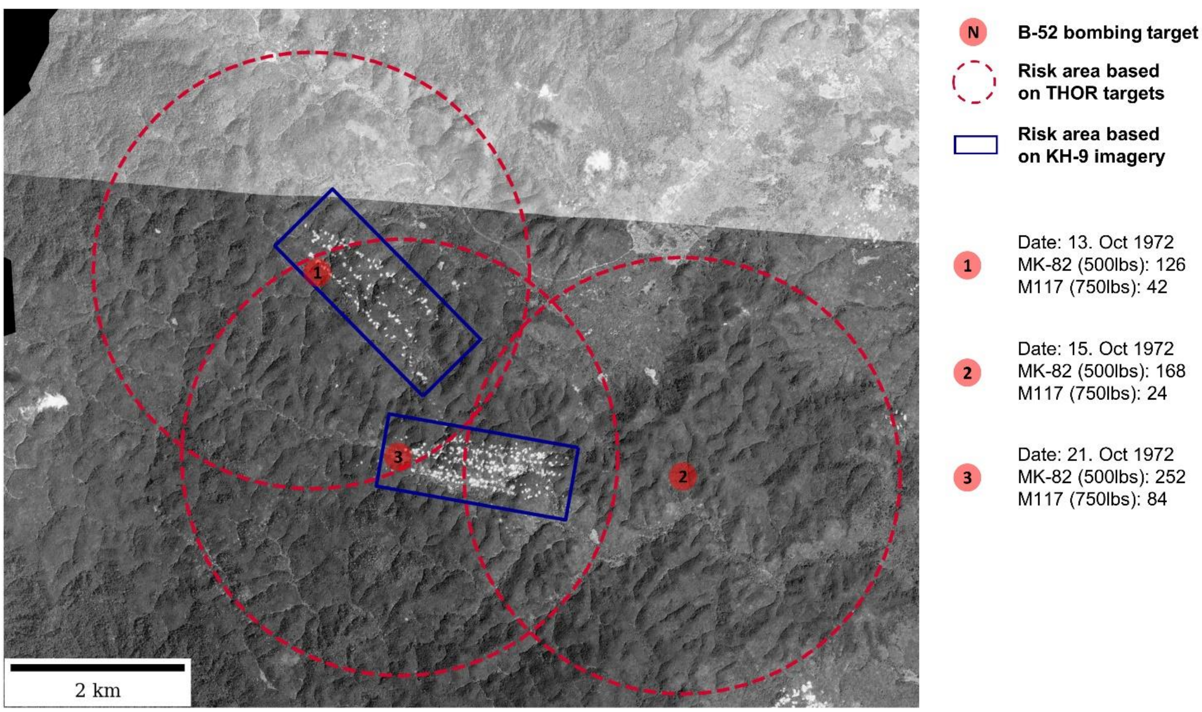


Figure 8. KH-9 imagery for an area in Kon Tum province showing three target locations of B-52 bombing missions that occurred during the month preceding the image acquisition. Overlayed on the imagery are estimated risk areas, delineating areas where unexploded bombs resulting from the bombing strikes could be located. These risk zones were determined by a 2.5 km radius around the THOR target locations (red) and rectangles drawn around the visible impact craters (blue).

Even in cases where bombings occurred near the time of image acquisition, not every dropped bomb recorded in THOR resulted in a crater detected by our model. In areas in QT province where bombing almost exclusively happened in the year before image acquisition, our model detected 150,895 craters, equivalent to 46% of bombs dropped that year. Several factors contribute to the lower number of detected craters, including: (1) bombs that left no craters, either because they exploded on water or failed to explode altogether; (2) craters that initially formed but vanished within less than a year due to human activities, natural events like landslides or consecutive bombing of the same location; (3) craters that were obscured in the imagery by clouds, vegetation, or flooding; and (4) craters that were visible in the imagery but not detected by our models.

While some of these limitations can be addressed, many are inherent to the approach. However, their impacts can be mitigated if they are recognized and dealt with correctly. Typically, bombing strikes involved dropping numerous bombs on a single target, and identifying half of the resulting craters can provide a sufficiently accurate representation of the affected area. The main challenge lies in recognizing and compensating for factors that introduce bias, such as crater visibility and model performance variations across different soil and land cover types. Further research is needed to investigate these factors and should incorporate multiple data sources including the THOR bombing data, historical land cover maps and confirmed locations of UXO.

3.3. Implications for mine action

The use of KH-9 imagery and derived crater locations could offer significant advantages to the mine action sector in Southeast Asia, extending beyond the capabilities of existing data sources used for non-technical surveys. Notably, our analysis revealed shortcomings of the THOR bombing data, emphasizing its lower precision compared to the detected crater locations. Additionally, the THOR data excludes weapons used by ground forces on both sides, such as artillery projectiles. Reports from local population carry a subjective element and are susceptible to recall bias, particularly when recounting events that happened 50 years ago. Additionally, their utility may be limited in previously unpopulated areas or where significant population shifts have happened since the war. Similarly, visual observations of UXO are invariably biased towards more populated areas. In contrast, the KH-9 imagery offers a more objective perspective, presenting an opportunity to address and overcome some of these challenges.

Despite the discussed benefits, the KH-9 imagery comes with its own biases and limitations. Due to their danger, mine action in Southeast Asia focuses on contamination with cluster submunitions which are only about the size of a tennis ball (McCosker et al., 2020). While patterns of smaller craters, that might be linked to artillery fire or cluster bomb strikes, were visible in certain areas of the imagery, these craters would have been too small to be detected by our models. However, even where impact craters are not directly visible in the KH-9 imagery, the presence of other objects, such as larger craters or military infrastructure, could be indicators for the presence of cluster submunitions. More research is needed to explore this possibility and should make use of existing clearance data. Additionally, despite the current focus on cluster submunitions, there are increasing efforts to understand and manage the residual risk from other weapon types (Stauffer & Mestre, 2020). The number of craters, as detected by our models, could be a valuable indicator to help determine the residual risk level for an area at a more fine-grained level than would be possible using only the bombing records.

One of the key strengths of the KH-9 imagery lies in its cost-effectiveness and ease of integration into existing workflows. Each image, covering a large area, only costs $30 on first request and previously requested images are freely available. The main limitation is the additional processing needed to orthorectify the images, including the time-consuming manual creation of ground control points. However, as demonstrated in our research, open-source tools can be used for this processing which reduces the cost. Products derived from our analysis can easily be integrated into existing mine action tools through imagery base layers for the KH-9 imagery and risk maps derived from detected bomb craters. The availability of the imagery for large parts of Southeast Asia makes it a useful tool for detailed analysis at both large (Figure 7) and small (Figure 8) scales.

3.4. Implications for sustainable development

Our work is directly aligned with Goal 16.1 of the Sustainable Development Goals (SDGs), which aims to "significantly reduce all forms of violence and related death rates everywhere". Additionally, mine action has been shown to have a direct impact on 12 out of the 17 SDGs (Hofmann & Juergensen, 2017). Notably, Lao PDR and Cambodia went as far as introducing an 18th SDG that specifically addresses the legacy of unexploded ordnance. The craters detected in this study allow for a detailed analysis of the impact of bombing on post-conflict land-use changes, which have previously been linked to deforestation (SDG 13, SDG 15), reduced agricultural productivity (SDG 2) and hindered infrastructure development (SDG 1, SDG 9, SDG

11) (Clerici et al., 2020; E. Lin, 2022; Martin et al., 2019; Munteanu et al., 2024; Ounmany & Andriesse, 2018).

In addition to supporting mine action, our work extends to other domains. While bomb craters have been identified as biodiversity hotspots (SDG 15) (Vad et al., 2017), they could also present potential public health hazards (SDG 3), as the stagnant water they collect can become breeding sites for mosquito larvae (Wimberly et al., 2021). Moreover, sediment buildup within these craters may contain concentrated levels of dioxins from herbicide spraying during the Vietnam War, posing a risk to individuals (SDG 3), particularly when the craters are repurposed as fish ponds (Olson & Morton, 2019). Bomb craters have been shown to alter hydrology and soil development in affected areas (Certini et al., 2013; Hupy & Koehler, 2012; Kiernan, 2015), but it remains unclear whether this could relate to the prevalence of landslides and flooding (SDG 13, SDG 15). More research is needed to understand these effects, and the bomb crater locations identified in this study could serve as a valuable resource for such investigations.

4 Conclusions

The presence of UXO in Vietnam, Lao PDR, and Cambodia continues to pose a significant threat to both public health and economic development. However, due to the expense and time required for detailed surveys, the exact locations of UXO often remain unknown. This study developed a workflow to orthorectify and automatically detect bomb craters in the declassified KH-9 imagery. The models achieved an overall F1-Score of 0.61 and predicted more than 500,000 bomb craters across the two study areas. The results demonstrate how the identified bomb craters can complement existing data sources such as the THOR bombing records. We estimate this could allow for more precise localization of suspected hazardous areas during non-technical surveys as well as a more fine-grained determination of residual risk of UXO in areas where extensive clearance operations are deemed too expensive. The developed methods are scalable to large regions at low cost and directly transferable to other affected areas in Southeast Asia. The instance segmentation workflow for the crater detection is also applicable to more recent conflicts including the ongoing war in Ukraine.

Acknowledgments

This work was supported by NERC through a SENSE CDT studentship (NE/T00939X/1) and by the Conflict and Environment Observatory. We thank the United States Geological Survey for providing the KH-9 imagery and the United States Department of Defense for providing the THOR bombing data. We thank Tim Horner for his valuable insights into mine action in Vietnam and Kinh Bac Dang for his support in identifying ground features in the KH-9 imagery. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (http://www.ecdf.ed.ac.uk/). During the preparation of this work the authors used ChatGPT (GPT-3.5) in order to improve the readability of the paper and to support the writing and documentation of code. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Open Research

The code used for the analysis in this study is available via https://github.com/pbarthelme/detecting-vietnam-war-bomb-craters and archived at https://doi.org/10.5281/zenodo.10709375 (Barthelme et al., 2024a). The predicted bomb craters and trained models are available at https://doi.org/10.5281/zenodo.10629987 (Barthelme et al., 2024b). The georeferenced KH-9 images are available via https://doi.org/10.7488/ds/7682 for Quang Tri province (Barthelme et al., 2024c) and via https://doi.org/10.7488/ds/7683 for the tri-border area (Barthelme et al., 2024d). The scanned KH-9 images (not georeferenced) are available as part of the Declassified Satellite Imagery - 3 collection courtesy of the U.S. Geological Survey (USGS EROS Center, 2018). They can be accessed via the EarthExplorer website https://earthexplorer.usgs.gov/ as part of the Data Set: Declassified Data > Declass 3 collection (Entity IDs for filtering the specific images used in the study are provided in the Supplementary Materials), download requires setting up a free account with USGS EROS web services. The THOR bombing records are available at https://data.world/datamil/vietnam-war-thor-data (File name: thor_data_vietnam.csv, last accessed 5. September 2023), download requires setting up a free account with data.world. Version 4.1 of the GADM administrative units used for creating some of the figures in this study are freely available for academic and other non-commerical use at www.gadm.org (last accessed: 2. February 2024). The SRTM GL1 dataset used for the orthorectification of the KH-9 imagery is available at OpenTopography via https://doi.org/10.5069/G9445JDF (NASA Shuttle Radar Topography Mission (SRTM), 2013). Version 3.0.0 of the Nasa Ames Stereo Pipeline used for orthorectification of the KH-9 imagery is preserved at https://doi.org/10.5281/ZENODO.5140581, available via Apache License 2.0 and developed openly at https://github.com/NeoGeographyToolkit/StereoPipeline (Beyer et al., 2021). Version 3.16.9 of QGIS used for the creation of ground control points is preserved at https://download.qgis.org/downloads/qgis-3.16.9.tar.bz2 (last accessed 2. Feburary 2024), available via GNU-General-Public-License and developed openly at https://github.com/qgis/QGIS (QGIS Association, 2023). The U-Net was implemented using the Python packages pytorch-cuda v11.7 (Paszke et al., 2019) and segmentation-models-pytorch v0.3.3 (Iakubovskii, 2019).

References

Anderson, D. L. (2002). The Vietnam War from A to Z. In *The Columbia Guide to the Vietnam War* (pp. 92–92). Columbia University Press. Retrieved from http://www.jstor.org/stable/10.7312/ande11492.17

Barthelme, P., Darbyshire, E., Spracklen, D., & Watmough, G. (2024b). Detecting Vietnam War bomb craters in declassified historical KH-9 satellite imagery - Data (Version 1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10629987

Barthelme, P., Darbyshire, E., Spracklen, D., & Watmough, G. (2024c). Georeferenced KH-9 Imagery - Quang Tri [Data set]. University of Edinburgh. School of GeoSciences. https://doi.org/10.7488/ds/7682

Barthelme, P., Darbyshire, E., Spracklen, D., & Watmough, G. (2024d). Georeferenced KH-9 Imagery - Tri-border area [Data set]. University of Edinburgh. School of GeoSciences. https://doi.org/10.7488/ds/7683

Barthelme, P., Darbyshire, E., Spracklen, D., & Watmough, G. (2024a). pbarthelme/detecting-vietnam-war-bomb-craters: Initial Release (Version 1.0.0) [Software]. Zenodo. https://doi.org/10.5281/zenodo.10709375

Bennett, M. M., Van Den Hoek, J., Zhao, B., & Prishchepov, A. V. (2022). Improving Satellite Monitoring of Armed Conflicts. *Earth's Future*, *10*(9), e2022EF002904. https://doi.org/10.1029/2022EF002904

Beyer, R., Alexandrov, O., ScottMcMichael, Broxton, M., Lundy, M., Husmann, K., et al. (2021). NeoGeographyToolkit/StereoPipeline 3.0.0 (Version 3.0.0) [Software]. Zenodo. https://doi.org/10.5281/zenodo.5140581

Bold, M., & Avenell, D. (2021). Data-Driven Decision-Making in Southeast Asia. *The Journal of Conventional Weapons Destruction*, *25*(1). Retrieved from https://commons.lib.jmu.edu/cisr-journal/vol25/iss1/18

Caicedo, J. C., Roth, J., Goodman, A., Becker, T., Karhohs, K. W., Broisin, M., et al. (2019). Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A*, *95*(9), 952–965. https://doi.org/10.1002/cyto.a.23863

Certini, G., Scalenghe, R., & Woods, W. I. (2013). The impact of warfare on the soil environment. *Earth-Science Reviews*, *127*, 1–15. https://doi.org/10.1016/j.earscirev.2013.08.009

Chen, D., Hu, F., Mathiopoulos, P. T., Zhang, Z., & Peethambaran, J. (2023). MC-UNet: Martian Crater Segmentation at Semantic and Instance Levels Using U-Net-Based Convolutional Neural Network. *Remote Sensing*, *15*(1). https://doi.org/10.3390/rs15010266

Clerici, N., Armenteras, D., Kareiva, P., Botero, R., Ramírez-Delgado, J. P., Forero-Medina, G., et al. (2020). Deforestation in Colombian protected areas increased during post-conflict periods. *Scientific Reports*, *10*(1), 4971. https://doi.org/10.1038/s41598-020-61861-y

Clermont, D., Kruse, C., Rottensteiner, F., & Heipke, C. (2019). Supervised detection of bomb craters in historical aerial images using convolutional neural networks. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(2), 67–74. https://doi.org/10.5194/isprs-archives-XLII-2-W16-67-2019

Cluster Munition Coalition. (2023). *Cluster Munition Monitor 2023*. (ICBL-CMC: August 2023). Retrieved from http://www.the-monitor.org/media/3383234/Cluster-Munition-Monitor-2023_Web.pdf

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–255). https://doi.org/10.1109/CVPR.2009.5206848

Deshpande, P., Belwalkar, A., Dikshit, O., & Tripathi, S. (2021). Historical land cover classification from CORONA imagery using convolutional neural networks and geometric moments. *International Journal of Remote Sensing*, *42*(13), 5144–5171. https://doi.org/10.1080/01431161.2021.1910365

Duncan, E. C., Skakun, S., Kariryaa, A., & Prishchepov, A. V. (2023). Detection and mapping of artillery craters with very high spatial resolution satellite imagery and deep learning. *Science of Remote Sensing*, *7*, 100092–100092. https://doi.org/10.1016/j.srs.2023.100092

Frost, A., Boyle, P., Autier, P., King, C., Zwijnenburg, W., Hewitson, D., & Sullivan, R. (2017). The effect of explosive remnants of war on global public health: a systematic mixed-studies review using narrative synthesis. *The Lancet Public Health*, *2*(6), e286–e296. https://doi.org/10.1016/S2468-2667(17)30099-3

Ghuffar, S., King, O., Guillet, G., Rupnik, E., & Bolch, T. (2023). Brief communication: Glacier mapping and change estimation using very high-resolution declassified Hexagon KH-9 panoramic stereo imagery (1971–1984). *The Cryosphere*, *17*(3), 1299–1306. https://doi.org/10.5194/tc-17-1299-2023

Gut, D., Tabor, Z., Szymkowski, M., Rozynek, M., Kucybała, I., & Wojciechowski, W. (2022). Benchmarking of Deep Architectures for Segmentation of Medical Images. *IEEE Transactions on Medical Imaging*, *41*(11), 3231–3241. https://doi.org/10.1109/TMI.2022.3180435

Hammer, E., FitzPatrick, M., & Ur, J. (2022). Succeeding CORONA: declassified HEXAGON intelligence imagery for archaeological and historical research. *Antiquity*, *96*(387), 679–695. https://doi.org/10.15184/aqy.2022.22

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *ArXiv Preprint, arXiv:1512.03385*. https://doi.org/10.48550/arXiv.1512.03385

High, H., Curran, J. R., & Robinson, G. (2013). Electronic Records of the Air War Over Southeast Asia: A Database Analysis. *Journal of Vietnamese Studies*, *8*(4), 86–124. https://doi.org/10.1525/vs.2014.8.4.86

Hofmann, U., & Juergensen, O. (2017). *Mine Action and the Sustainable Development Goals*. Geneva International Centre for Humanitarian Demining & United Nations Development Programme. Retrieved from https://www.undp.org/publications/mine-action-and-sustainable-development-goals

Hupy, J. P., & Koehler, T. (2012). Modern warfare as a significant form of zoogeomorphic disturbance upon the landscape. *Geomorphology*, *157–158*, 169–182. https://doi.org/10.1016/j.geomorph.2011.05.024

Iakubovskii, P. (2019). Segmentation Models Pytorch (Version 0.3.3) [Software]. GitHub. Retrieved from https://github.com/qubvel/segmentation_models.pytorch

Kiernan, K. (2015). Nature, Severity and Persistence of Geomorphological Damage Caused by Armed Conflict. *Land Degradation & Development*, *26*(4), 380–396. https://doi.org/10.1002/ldr.2216

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv Preprint ArXiv:1412.6980*. https://doi.org/10.48550/arXiv.1412.6980

Kruse, C., Rottensteiner, F., & Heipke, C. (2019). Marked Point Processes for the Automatic Detection of Bomb Craters in Aerial Wartime Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-2-W13*, 51–60. https://doi.org/10.5194/isprs-archives-XLII-2-W13-51-2019

Kugelman, J., Allman, J., Read, S. A., Vincent, S. J., Tong, J., Kalloniatis, M., et al. (2022). A comparison of deep learning U-Net architectures for posterior segment OCT retinal layer segmentation. *Scientific Reports*, *12*(1), 14888. https://doi.org/10.1038/s41598-022-18646-2

Lasaponara, R., Yang, R., Chen, F., Li, X., & Masini, N. (2018). Corona Satellite Pictures for Archaeological Studies: A Review and Application to the Lost Forbidden City of the Han–Wei Dynasties. *Surveys in Geophysics 2018 39:6*, *39*(6), 1303–1322. https://doi.org/10.1007/S10712-018-9490-2

Le, D. T., Pham, T. M., & Polachek, S. (2022). The long-term health impact of Agent Orange: Evidence from the Vietnam War. *World Development*, *155*, 105813–105813. https://doi.org/10.1016/J.WORLDDEV.2022.105813

Le, K., & Nguyen, M. (2020). Aerial bombardment and educational attainment. *International Review of Applied Economics*, *34*(3), 361–383. https://doi.org/10.1080/02692171.2020.1736012

Lin, E. (2022). How War Changes Land: Soil Fertility, Unexploded Bombs, and the Underdevelopment of Cambodia. *American Journal of Political Science*, *66*(1), 222–237. https://doi.org/10.1111/ajps.12577

Lin, E., Qin, R., Edgerton, J., & Kong, D. (2020). Crater detection from commercial satellite imagery to estimate unexploded ordnance in Cambodian agricultural land. *PLOS ONE*, *15*(3), e0229826. https://doi.org/10.1371/journal.pone.0229826

Lin, T.-Y., Goyal, P., Girshick, R. B., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *ArXiv Preprint*, *arXiv:1708.02002*. https://doi.org/10.48550/arXiv.1708.02002

Martin, M. F., Dolven, B., Feickert, A., & Lum, T. (2019). War Legacy Issues in Southeast Asia: Unexploded Ordnance (UXO). *Current Politics and Economics of South, Southeastern, and Central Asia*, *28*(2/3), 199–230.

McCosker, K., Stoa, J., & Harrison, K. (2020). Cluster Munition Remnant Survey in Southeast Asia. *The Journal of Conventional Weapons Destruction*, *23*(3). Retrieved from https://commons.lib.jmu.edu/cisr-journal/vol23/iss3/6

Miguel, E., & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, *96*(1), 1–15. https://doi.org/10.1016/j.jdeveco.2010.07.004

Mulyanto, M., Faisal, M., Prakosa, S. W., & Leu, J.-S. (2021). Effectiveness of Focal Loss for Minority Classification in Network Intrusion Detection Systems. *Symmetry*, *13*(1). https://doi.org/10.3390/sym13010004

Munteanu, C., Kraemer, B. M., Hansen, H. H., Miguel, S., Milner-Gulland, E. J., Nita, M., et al. (2024). The potential of historical spy-satellite imagery to support research in ecology and conservation. *BioScience*, biae002. https://doi.org/10.1093/biosci/biae002

NASA Shuttle Radar Topography Mission (SRTM). (2013). Shuttle Radar Topography Mission (SRTM) Global [Data set]. Distributed by OpenTopography. https://doi.org/10.5069/G9445JDF

Nguyen, L. B. (2020). Land cover change detection in northwestern Vietnam using Landsat images and Google Earth Engine. *Journal of Water and Land Development*, *no. 46*(46), 162–169. https://doi.org/10.24425/JWLD.2020.134209

Nita, M. D., Munteanu, C., Gutman, G., Abrudan, I. V., & Radeloff, V. C. (2018). Widespread forest cutting in the aftermath of World War II captured by broad-scale historical Corona spy satellite photography. *Remote Sensing of Environment*, *204*, 322–332. https://doi.org/10.1016/J.RSE.2017.10.021

Olson, K. R., & Morton, L. W. (2019). Long-Term Fate of Agent Orange and Dioxin TCDD Contaminated Soils and Sediments in Vietnam Hotspots. *Open Journal of Soil Science*, *09*(01), 1. https://doi.org/10.4236/ojss.2019.91001

Ounmany, K., & Andriesse, E. (2018). The Legacy of the Vietnam War: Making a Living amid Unexploded Ordnance in Xieng Khouang Province, Northern Laos. *Asian Studies Review*, *42*(3), 439–458. https://doi.org/10.1080/10357823.2018.1481365

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *ArXiv Preprint*, *arXiv:1912.01703*. https://doi.org/10.48550/arXiv.1912.01703

QGIS Association. (2023). QGIS Geographic Information System (Version 3.16.9) [Software]. Retrieved from http://www.qgis.org/

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia and Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Silburt, A., Ali-Dib, M., Zhu, C., Jackson, A., Valencia, D., Kissin, Y., et al. (2019). Lunar crater identification via deep learning. *Icarus*, *317*, 27–38. https://doi.org/10.1016/j.icarus.2018.06.022

Sohn, H.-G., Kim, G.-H., & Yom, J.-H. (2004). Mathematical modelling of historical reconnaissance CORONA KH-4B Imagery. *The Photogrammetric Record*, *19*(105), 51–66. https://doi.org/10.1046/j.0031-868X.2003.00257.x

Stauffer, K., & Mestre, C. (2020). Long-Term Risk Management Tools for Protocols for Residual Explosive Ordnance Mitigation: A Pretest in Vietnam. *The Journal of Conventional Weapons Destruction*, *23*(3). Retrieved from https://commons.lib.jmu.edu/cisr-journal/vol23/iss3/5

UNMAS. (2019). IMAS 08.10 Non-technical survey. Retrieved February 5, 2024, from https://www.mineactionstandards.org/standards/08-10/

USGS EROS Center. (2018). Declassified Satellite Imagery - 3 [Data set]. https://doi.org/10.5066/F7WD3Z10

Vad, C. F., Péntek, A. L., Cozma, N. J., Földi, A., Tóth, A., Tóth, B., et al. (2017). Wartime scars or reservoirs of biodiversity? The value of bomb crater ponds in aquatic conservation. *Biological Conservation*, *209*, 253–262. https://doi.org/10.1016/j.biocon.2017.02.025

Waga, J. M., Szypuła, B., & Fajer, M. (2022). The Archaeology of Unexploded World War II Bomb Sites in the Koźle Basin, Southern Poland. *International Journal of Historical Archaeology*. https://doi.org/10.1007/s10761-022-00672-5

Wang, H., & Miao, F. (2022). Building extraction from remote sensing images using deep residual U-Net. *European Journal of Remote Sensing*, *55*(1), 71–85. https://doi.org/10.1080/22797254.2021.2018944

Wimberly, M. C., Beurs, K. M. de, Loboda, T. V., & Pan, W. K. (2021). Satellite Observations and Malaria: New Opportunities for Research and Applications. *Trends in Parasitology*, *37*(6), 525–537. https://doi.org/10.1016/j.pt.2021.03.003

Yamada, T., & Yamada, H. (2021). The long-term causal effect of U.S. bombing missions on economic development: Evidence from the Ho Chi Minh Trail and Xieng Khouang Province in Lao P.D.R. *Journal of Development Economics*, *150*, 102611–102611. https://doi.org/10.1016/J.JDEVECO.2020.102611

Zhou, Y., Chen, G., Qiao, X., & Lu, L. (2021). Mining High-Resolution KH-9 Panoramic Imagery to Determine Earthquake Deformation: Methods and Applications. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2021.3116441

References From the Supporting Information

Aschauer, J., & Kenkmann, T. (2017). Impact cratering on slopes. *Icarus*, *290*, 89–95. https://doi.org/10.1016/j.icarus.2017.02.021

Barlow, N. G., Stopar, J. D., & Hargitai, H. (2021). Ejecta (Impact). In H. Hargitai & Á. Kereszturi (Eds.), *Encyclopedia of Planetary Landforms* (pp. 1–8). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-9213-9_114-1

Hayashi, K., & Sumita, I. (2017). Low-velocity impact cratering experiments in granular slopes. *Icarus*, *291*, 160–175. https://doi.org/10.1016/j.icarus.2017.03.027

Roberts, A. L., Fawdon, P., & Mirino, M. (2021). Impact crater degradation, Oxia Planum, Mars. *Journal of Maps*, *17*(2), 581–590. https://doi.org/10.1080/17445647.2021.1976685

Sabuwala, T., Butcher, C., Gioia, G., & Chakraborty, P. (2018). Ray Systems in Granular Cratering. *Physical Review Letters*, *120*(26), 264501–264501. https://doi.org/10.1103/PhysRevLett.120.264501

Earth's Future

Supporting Information for

Detecting Vietnam War Bomb Craters in Declassified Historical KH-9 Satellite Imagery

Philipp Barthelme[1], Eoghan Darbyshire[2], Dominick Spracklen[3], and Gary R Watmough[1,4,5]

[1] School of Geosciences, University of Edinburgh, Edinburgh, United Kingdom.

[2] Conflict and Environment Observatory, Mytholmroyd, United Kingdom.

[3] School of Earth and Environment, University of Leeds, Leeds, United Kingdom.

[4] Global Academy of Agriculture and Food Systems, University of Edinburgh, United Kingdom

[5] Novel Data Ecosystems for Sustainability (NODES) Research Group. Advancing Systems Analysis Program, International Institute for Applied Systems Analysis (IIASA), Austria

Contents of this file

Introduction

The following gives additional information about the methods used for crater labeling (Text S1) and the processing of the THOR bombing records (Text S2, Figure S1, Figure S2, Table S2, Table S3). To assess the robustness of the results to the THOR processing, Tables S4 and S5 provide spearman correlations between detected bomb craters and THOR bombing for different subsets of THOR, extending Table 2 in the paper. In addition, Table S1 lists the KH-9 images used during the analysis.

Text S1. Crater labeling

Simple craters usually consist of a bowl-shaped hole with an elevated rim and a circular continuous ejecta blanket around the rim (Barlow et al., 2021; Roberts et al., 2021). In the KH-9 imagery, crater bowl and rim were usually visible as a bright circle with a shadow on the side of the crater that faces the sun, or as a dark circle if the crater was filled with water. The crater ejecta were often visible as a bright circle around the crater bowl and in some cases as rayed ejecta extending much further from the crater (Sabuwala et al., 2018). We labeled an object as a crater of type *Rim* if both the crater bowl and the crater ejecta were visible and could be distinguished from each other.

In the imagery, small craters were often only visible as circular white blobs, where the crater bowl could not be distinguished from the crater ejecta. We labeled these circular white blobs as craters of type *Pattern* if they appeared in patterns with other blobs of the same size. We also ensured that the context did not indicate them to be different objects such as houses, trees or circular graves. This was done using both the context visible in the KH-9 imagery and, if necessary, current Google Earth imagery. As the image resolution and quality made it difficult to reliably identify very small objects as craters, we excluded all objects smaller than 25 pixels, equivalent to 25 m^2. Similar to our approach, Lin et al. (2020) limit their analysis to bomb craters with diameters between 3 and 12 m using satellite imagery with a resolution of 0.5 m. Duncan et al. (2023), who also use imagery with a resolution of 0.5 m, do not set a size limit during crater labeling but find that their model performs worse on smaller craters of less than 30 m^2. They also note the difficulty of labeling craters on vegetated and heterogeneous surfaces and limit their study to agricultural fields with short vegetation.

We labeled craters as type *Group* if crater bowls or continuous ejecta of three or more craters overlapped, in which case only the crater bowl was labeled as the crater ejecta could not be attributed clearly to an individual crater. We labeled craters as type *Bowl* if only the crater bowl, often filled with water, was visible. This can occur for older craters where the ejecta has eroded over time. We labeled craters as type *Crescent* where we saw a crescent shaped crater rim/ejecta, which can occur in areas with steep slopes or where erosion has affected the crater appearance (Aschauer & Kenkmann, 2017; Hayashi & Sumita, 2017). Craters of types *Group*, *Bowl* and *Crescent* were difficult to identify reliably, and we often relied on context to make the final decision.

We used the circle shape from the QGIS toolbar to label the craters, even if they were not perfectly circular, as we found this to be an efficient approach. We used freehand label shapes to separate overlapping crater bowls. The labels encompassed both the crater bowl and, if visible, the crater ejecta, but excluded rayed ejecta. For craters of type *Group*, only the crater bowl was labeled since the continuous ejecta could not be attributed clearly to any individual crater. For craters of type *Crescent*, for which a full circular shape is not given, we still used a circle shape as the label, based on the best approximation of the suspected crater shape.

Text S2. THOR processing

This section provides additional information about issues with the THOR data that we identified during our analysis and describes how we decided to correct or circumvent these issues. This is not an exhaustive list of issues in the THOR data but focuses on the issues that were immediately relevant to our analysis. We chose solutions that were acceptable for the purpose of our research, which compares the bombing on a large scale. However, some individual bombing records might not be processed correctly, which is unavoidable due to the limitations of the THOR data. More work is needed to identify additional limitations, but this was outside the scope of this paper.

1. *Coordinate reference system*

Our analysis suggested that the THOR target coordinates, provided in columns *tgtlatdd_ddd_wgs84* and *tgtlonddd_ddd_wgs84*, despite their names, did not use the WGS84 coordinate reference system but were instead provided in the Indian 1960 geodetic coordinate system (EPSG:4131). We therefore converted these coordinates from EPSG:4131 to EPSG:4326 (using EPSG:1542), which led to a shift of about 500 meters. We found that the resulting coordinates matched the locations of the craters visible in the KH-9 imagery more closely. This was particularly true for B-52 bombing missions where it was often easiest to match the distinct lines of craters to individual records in the THOR data. We also tested the hypothesis using historical topographic maps[1] created by the U.S. Army Map Service during the Vietnam War. These maps show the Military Grid Reference System (MGRS) coordinates, which were used by the U.S. military at the time. While the THOR bombing records do not contain the original MGRS target coordinates, these coordinates are included in the corresponding SEADAB source records available in the National Archives and Records Administration (NARA)

[1] These maps are part of the *AMS Topographic Maps - Series L7014* available at https://maps.lib.utexas.edu/maps/topo/vietnam/ (accessed 21/10/2023)

archives[2]. We again found that the shifted target coordinates matched the MGRS target coordinates more closely when we located those on the topographic maps.

2. Mission function codes

We found that multiple mission function descriptions, given in the *mfunc_desc* column in the THOR data, were wrong which, in some cases, also led to a wrong mission function class (*mfunc_desc_class*). We suspected that during the conversion from mission function codes (*mfunc*) to mission function descriptions, the CACTA lookup table was used for both CACTA and SEADAB records. However, some of the code meanings were changed for the SEADAB data, which can be seen in the corresponding NARA documentation[3]. Most notably, B-52 bombing missions were wrongly classified in the THOR data as non-kinetic missions with description "COMBT CARGO AIR DROP" instead of the correct "HEAVY BOMBARD". The wrong mission functions can be corrected by using the correct lookup table, as provided in the NARA documentation for the SEADAB data, to map the mission function codes (*mfunc*) to the mission function description (*mfunc_desc*). However, we did not use the mission function information in our analysis and therefore did not apply this processing step. We nevertheless make note of this information here as filtering the kinetic records in THOR by using the mission function class has been a common processing step in past analyses. We refer the reader to the corresponding documentation of the SEADAB and CACTA data in the NARA archives for the correct mission function mapping.

3. B-52 bombing records

We found that B-52 bombing missions were likely recorded in both the SEADAB and the SACCOACT databases starting from March, 1 1971. Figure S1 shows the number of weapons dropped in B-52 bombing missions by source database. While the SACCOACT and SEADAB records do not match exactly, they approximately agree from early 1971 onwards. We confirmed this by checking individual lines of craters in the KH-9 imagery that often can be matched to both a SEADAB and a SACCOACT B-52 bombing record despite only showing craters for one bomb strike. Notably, the target coordinates from the SACCOACT records

[2] The NARA SEADAB records (National Archives Identifier (NAID): 602566) are available online at https://catalog.archives.gov/id/602566 (accessed 21/10/2023)

[3] p.84-85 in the technical documentation for the NARA SEADAB data (NAID: 602566) https://s3.amazonaws.com/NARAprodstorage/opastorage/live/92/9370/1937092/content/arcmedia/electronic-records/rg-218/seadab/123.1DP.pdf (accessed 21/10/2023) and p.95-96 in the technical documentation of the CACTA data (NAID: 602566) https://s3.amazonaws.com/NARAprodstorage/opastorage/live/45/5547/2554745/content/electronic-records/rg-218/CACTA/136.1DP.pdf (accessed 21/10/2023)

did not exactly match the SEADAB target coordinates. We found that the SACCOACT targets were much less precise and often appeared to be located on a grid of about one nautical mile (approx. 1.8 km). In contrast, the SEADAB targets appeared to be more precise as they more closely matched the lines of craters identified in the KH-9 imagery. This can also be seen in Figure S2, where we compared the spatial distribution of the B-52 bombing missions for the Quang Tri study area. While the numbers aggregated by grid cells (Figure S2a and b) matched well between the databases, Figure S2c and d clearly show the better precision of the SEADAB records. The SEADAB data also appeared to contain additional records for some of the time periods (see Figure S1). We therefore decided to keep the SEADAB records and discarded all SACCOACT records after March, 1 1971.

4. Missing weapon types in SEADAB

Many records in the THOR data, including about 1.4 million of the 1.8 million total records originating from the SEADAB database, did not contain information about the weapon type used. As the weapon type information was crucial for our analysis, we imputed some of the missing weapon types based on the estimated weight of the weapon type, which was calculated using the equation

$$weapontypeweight_{est} = \frac{1}{10} \times \frac{weaponsloadedweight}{numweaponsdelivered} \tag{1}$$

where *weaponsloadedweight* describes the weight loaded on the planes flying the mission and *numweaponsdelivered* denotes the total number of weapons loaded on the plane. The imputed weapon types and their corresponding estimated weight are provided in Table S1. However, the estimated weight of a weapon type sometimes matched with multiple potential weapon types. For the cases given in Table S2, this was not an issue as either all the matching weapon types would also be relevant for our analysis (see Section 5) or the number of records for other weapon types were very low. One exception was the estimated weapon type weight of 820 pounds, which matched both the “M117 GP BOMB (750) LD” as well as the “CBU49 AN PR MINE”. To allow for a more accurate matching, we took the type of plane into account as we found that the “M117 GP BOMB (750) LD” bombs were more likely to be dropped from one of the following planes: "A-1", "A-37", "B-52", "B-57", "F-100", "F-105" and "F-5". In cases where a different plane was used, we matched the estimated weight of 820 pounds to the “CBU49 AN PR MINE”.

5. Weapon types resulting in large craters

We selected 22 weapon types for our analysis and discarded all records with other weapon types. The decision was made based on our understanding of which weapon types would result in large craters detectable by our model ($\geq 25m^2$). We also removed some weapon types that might result in large craters but which had very few records associated with them as these did not substantially affect our results. Removing those records made it easier in cases where it was challenging to understand the weapon type from the given name or to impute the weapon type based on its weight. The list of weapon types used for the analysis is given in Table S3.

6. *Maximum of weapons per plane*

We found multiple records with large amounts of weapons dropped that far exceeded the number of weapons that could be carried by the corresponding aircraft. In some cases, this could be traced back to simple typos, but investigating every individual case would be time consuming and often it was unclear if a record could be adjusted or was completely wrong. The maximum number of large bombs an individual plane could carry during the war was 108 (by a B-52 bomber[4]), and as we only considered large bombs in our analysis, we removed all records where the number of weapons per plane exceeded 108. A more sophisticated way to address the issue would be to consider the maximum load of individual aircraft types, but this would be complicated due to the large number of combinations of weapons and aircraft types. Therefore, we opted for this simpler solution which was sufficient for the purpose of our analysis as it removes the most severe errors.

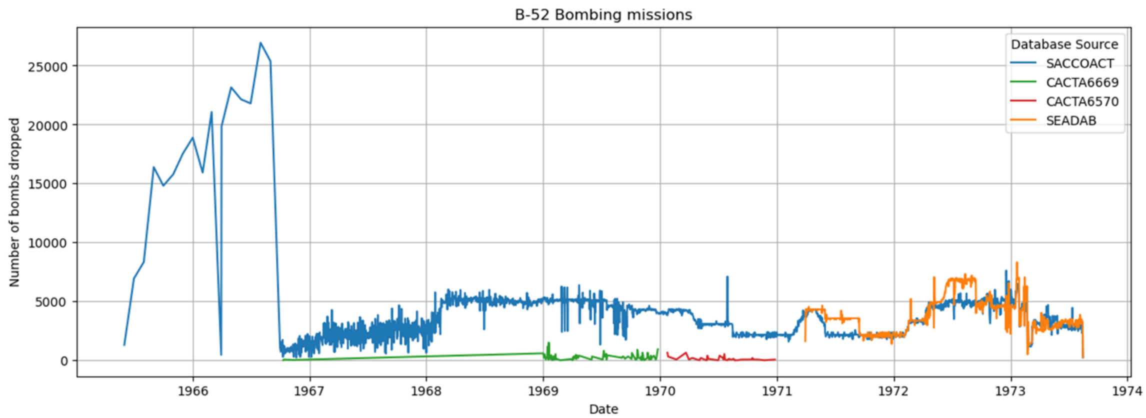
7. *Correlation results for different subsets of the THOR data*

We calculated the correlations between the detected bomb craters and different subsets of the THOR records, aggregated by 2×2 km grid cells, in order to test the robustness of the results we presented in the paper (see Tables S4 and S5). As expected, we typically saw higher correlations between craters and bombing records when only considering weapons dropped during the year before the KH-9 images were taken. We also saw higher correlations for the subset of heavier kinetic weapons compared to weapon types weighing less than 200 pounds. Filtering on non-kinetic weapons resulted in much lower correlations, which was expected as these weapon types would not result in craters. Correlations were higher for the B-52 missions recorded in SEADAB compared to the SACCOACT records when considering previous year bombings,

[4] This number corresponds to the B-52D model which was able to carry up to 84 bombs internally and an additional 24 bombs under its wings https://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact-Sheets/Display/Article/195838/the-big-belly-bomber/ (accessed 21/10/2023)

which we suspect was due to the lower precision of the SACCOACT target locations (see Section 3). Overall, the correlation results showed that our analysis was robust and confirmed some of the methods we used for processing the THOR data.

a. Daily number of bombs dropped on B-52 bombing missions recorded in THOR by source database



b. Daily number of bombs dropped on B-52 bombing missions recorded in THOR 1971- 1973

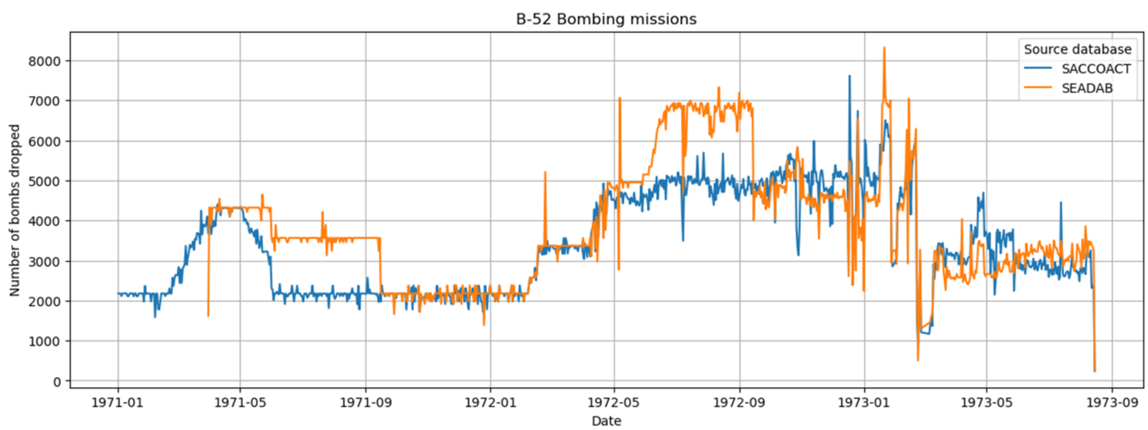


Figure S1. Daily number of bombs dropped on B-52 bombing missions by source database, with the full duration shown in (a) and the period after 1971 shown in (b).

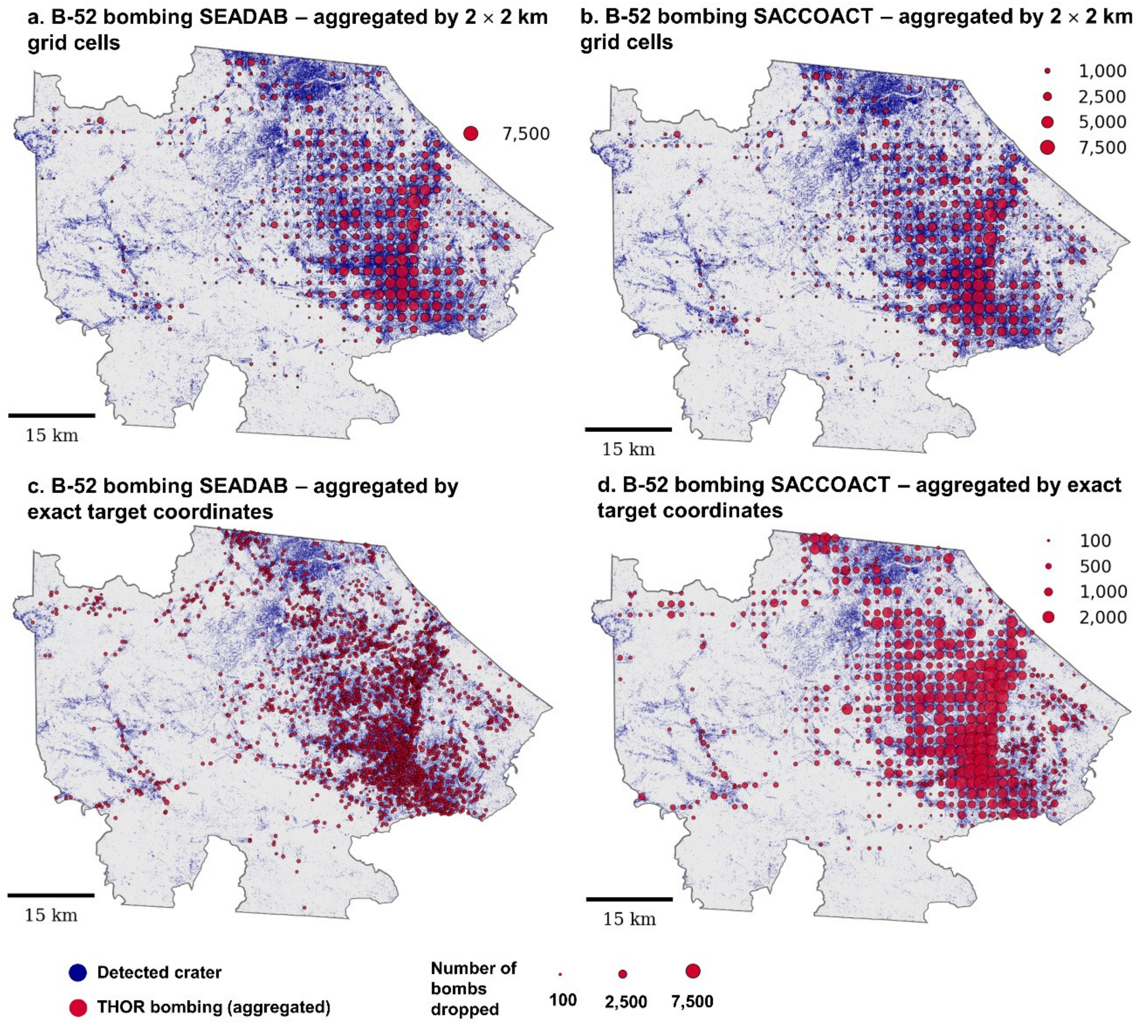


Figure S2. B-52 bombing mission during the year before the KH-9 images were taken (March 1972 – March 1973) split up by source database. Panels (a) and (b) show the number of bombs aggregated by a grid of 2 × 2 km whereas (c) and (d) show the number of bombs dropped aggregated by exact target location. The grid pattern visible in panel (d) likely arises from the less precise target locations in the SACCOACT database which lead to multiple records with the exact same target location, often located on a grid of about one nautical mile (~1.8km). However, panel (c) shows that more precise target locations were recorded at the time as they are available in the SEADAB database.

Entity ID	Study area	Acquisition data
D3C1205-100113A009	Quang Tri	20/03/1973
D3C1205-100113F009	Quang Tri	20/03/1973
D3C1205-100113A010	Quang Tri	20/03/1973
D3C1205-100113F010	Quang Tri	20/03/1973
D3C1205-100113A011	Quang Tri	20/03/1973
D3C1205-100113F011	Quang Tri	20/03/1973
D3C1205-100113A012	Quang Tri	20/03/1973
D3C1205-100113F012	Quang Tri	20/03/1973
D3C1204-200292A077	Tri-border area	04/11/1972
D3C1204-200292F077	Tri-border area	04/11/1972
D3C1204-200292A078	Tri-border area	04/11/1972
D3C1204-200292F078	Tri-border area	04/11/1972
D3C1204-200292A079	Tri-border area	04/11/1972
D3C1204-200292F079	Tri-border area	04/11/1972
D3C1204-200292A080	Tri-border area	04/11/1972
D3C1204-200292F080	Tri-border area	04/11/1972
D3C1204-200292A081	Tri-border area	04/11/1972
D3C1204-200292F081	Tri-border area	04/11/1972
D3C1204-200292A082	Tri-border area	04/11/1972
D3C1204-200292F082	Tri-border area	04/11/1972

Table S1. KH-9 images used in the analysis. Stereo pairs were used during orthorectification but only the aft looking images were used for the crater detection.

Estimated weapon type weight (in pounds)	Matched weapon type
260	MK81 GP BOMB (250)
531	MK 82 GP BOMB (500) LD
571	MK82 GP BOMB (500) HD
820	M117 GP BOMB (750) LD/ CBU49 AN PR MINE
1100	MK83 GP BOMB (1000)

Table S2. Matching between estimated weapon type weight and weapon type.

Weapon type	Source database
500LB GP MK-82	CACTA
750LB GP M-117	CACTA
250LB MK-81	CACTA
500LB GP M-64	CACTA
250LB M-57	CACTA
200/260 M81/88	CACTA
1000LB MK-83	CACTA
100LB GP M-30	CACTA
1000LB GP M-65	CACTA
2000LB MK-84	CACTA
2000LB M-66	CACTA
3000LB M-118	CACTA
MK 82 GP BOMB (500) LD	SEADAB
M117 GP BOMB (750) LD	SEADAB
MK81 GP BOMB (250)	SEADAB
MK82 GP BOMB (500) HD	SEADAB
MK 82 GP BOMB (500)	SEADAB
MK83 GP BOMB (1000)	SEADAB
MK82 B	SACCOACT
750 GP	SACCOACT
M64A1	SACCOACT
MK83 B	SACCOACT

Table S3. Weapon types and corresponding source database for the weapon types used in the analysis.

Category	Craters	Pattern	Rim	Group	Crescent	Bowl	Number of Weapons
Total bombing[5]	0.58	0.55	0.46	0.33	0.61	0.41	2,232,280
Bombing previous year	0.76	0.74	0.78	0.68	0.52	0.62	654,730
Kinetic weapons[6] over 200 pounds	0.58	0.55	0.46	0.34	0.62	0.42	2,364,961
Kinetic weapons over 200 pounds previous year	0.76	0.74	0.78	0.69	0.51	0.62	713,599
Kinetic weapons under 200 pounds	0.12	0.10	0.02	-0.07	0.29	0.10	316,442
Kinetic weapons under 200 pounds previous year	0.52	0.50	0.54	0.48	0.35	0.45	15,472
Non-kinetic weapons[7]	0.26	0.24	0.20	0.05	0.31	0.22	52,238
Non-kinetic weapons previous year	0.38	0.36	0.35	0.27	0.27	0.32	5,111
Unknown weapon type	0.76	0.74	0.71	0.61	0.60	0.58	702,797
Unknown weapon type previous year	0.77	0.74	0.79	0.69	0.52	0.64	489,716
B-52 SACCOACT previous year	0.72	0.70	0.73	0.68	0.48	0.54	438,411
B-52 SEADAB previous year	0.73	0.71	0.75	0.70	0.49	0.56	456,992
B-52 SACCOACT	0.53	0.50	0.44	0.36	0.51	0.36	1,210,553
B-52 SEADAB	0.68	0.69	0.61	0.53	0.56	0.44	687,490

Table S4. Spearman correlation coefficients between detected craters and number of weapons dropped according to the THOR bombing data aggregated across grid cells of 2 km × 2km in the Quang Tri study area. The number of weapons dropped based on the THOR data is provided for additional context. For the "Craters" category detected craters of all crater classes were aggregated before calculating the correlation. The descriptor

[5] Total bombing refers to the final processing of the THOR data for our analysis, as presented in the paper itself, and therefore only considers the weapon types in Table S2

[6] Kinetic weapons refer to the classification provided by the *mfunc_desc_class* column. Here we consider the B-52 bombings with mission function code 61 originating from the SEADAB data as kinetic despite them being classified as non-kinetic in THOR. However, we did not update any other wrongly mapped mission function codes (see Text S2 Part 2 for details). No other filters, such as selecting specific weapon types (Text S2 Part 5) or removing records with too many bombs per plane (Text S2 Part 6) were applied.

[7] Non-kinetic weapons refer to the classification provided by the *mfunc_desc_class* column. We consider the B-52 bombing missions originating from SEADAB as kinetic and therefore exclude them here. No other filters were applied. Case numbers are low as most non-kinetic missions record 0 for the *numweaponsdelivered* field.

“previous year” was added where only missions that took place during the year before the KH-9 images were taken were considered, otherwise all missions before the KH-9 images were taken were considered.

Category	Craters	Pattern	Rim	Group	Crescent	Bowl	Number of Weapons
Bombing	0.52	0.52	0.4	0.1	0.45	0.33	1,133,025
Bombing previous year	0.51	0.53	0.42	0.11	0.47	0.33	321,504
Kinetic weapons over 200 pounds	0.52	0.52	0.4	0.1	0.45	0.34	1,250,068
Kinetic weapons over 200 pounds previous year	0.51	0.52	0.41	0.12	0.46	0.32	343,853
Kinetic weapons under 200 pounds	0.41	0.4	0.32	0.09	0.35	0.29	534,798
Kinetic weapons under 200 pounds previous year	0.34	0.34	0.29	0.13	0.29	0.25	149,409
Non-kinetic weapons	0.17	0.18	0.13	0.07	0.14	0.1	78,338
Non-kinetic weapons previous year	0.12	0.13	0.1	0.08	0.11	0.07	3,011
Unknown weapon type	0.49	0.49	0.39	0.11	0.43	0.34	600,066
Unknown weapon type previous year	0.5	0.5	0.4	0.12	0.44	0.32	334,623
B-52 SACCOACT previous year	0.43	0.44	0.38	0.07	0.4	0.3	220,174
B-52 SEADAB previous year	0.46	0.47	0.38	0.1	0.43	0.31	239,741
B-52 SACCOACT	0.47	0.48	0.38	0.08	0.43	0.31	711,533
B-52 SEADAB	0.48	0.49	0.4	0.09	0.44	0.33	294,806

Table S5. Spearman correlation coefficients between detected craters and number of weapons dropped according to the THOR bombing data aggregated across grid cells of 2 km × 2km in the Tri-border study area. The number of weapons dropped based on the THOR data is provided for additional context. For the “Craters” category detected craters of all crater classes were aggregated before calculating the correlation. The descriptor “previous year” was added where only missions that took place during the year before the KH-9 images were taken were considered, otherwise all missions before the KH-9 images were taken were considered.

References:

Aschauer, J., & Kenkmann, T. (2017). Impact cratering on slopes. *Icarus*, *290*, 89–95. https://doi.org/10.1016/j.icarus.2017.02.021

Barlow, N. G., Stopar, J. D., & Hargitai, H. (2021). Ejecta (Impact). In H. Hargitai & Á. Kereszturi (Eds.), *Encyclopedia of Planetary Landforms* (pp. 1–8). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-9213-9_114-1

Duncan, E. C., Skakun, S., Kariryaa, A., & Prishchepov, A. V. (2023). Detection and mapping of artillery craters with very high spatial resolution satellite imagery and deep learning. *Science of Remote Sensing*, *7*, 100092–100092. https://doi.org/10.1016/j.srs.2023.100092

Hayashi, K., & Sumita, I. (2017). Low-velocity impact cratering experiments in granular slopes. *Icarus*, *291*, 160–175. https://doi.org/10.1016/j.icarus.2017.03.027

Lin, E., Qin, R., Edgerton, J., & Kong, D. (2020). Crater detection from commercial satellite imagery to estimate unexploded ordnance in Cambodian agricultural land. *PLOS ONE*, *15*(3), e0229826. https://doi.org/10.1371/journal.pone.0229826

Roberts, A. L., Fawdon, P., & Mirino, M. (2021). Impact crater degradation, Oxia Planum, Mars. *Journal of Maps*, *17*(2), 581–590. https://doi.org/10.1080/17445647.2021.1976685

Sabuwala, T., Butcher, C., Gioia, G., & Chakraborty, P. (2018). Ray Systems in Granular Cratering. *Physical Review Letters*, *120*(26), 264501–264501. https://doi.org/10.1103/PhysRevLett.120.264501